Least-Squares Methods for the Solution of Fluid-Structure Interaction Problems

Von der Carl-Friedrich-Gauß-Fakultät für Mathematik und Informatik der Technischen Universität Braunschweig

genehmigte Dissertation

zur Erlangung des Grades eines Doktor-Ingenieurs (Dr.-Ing.)

von
Oliver Kayser-Herold
aus Bremen

Prüfung am 26. Januar 2006

1. Referent: Prof. H.G. Matthies, Ph.D.
2. Referent: Prof. Dr. G. Starke
eingereicht am 18. Oktober 2005

Kurzfassung

Zur Lösung von partiellen Differentialgleichungen (PDE) sind verschiedene numerische Verfahren vorgeschlagen worden. Die meisten basieren auf einem Variationsprinzip, das die PDE in eine äquivalente Integralgleichung überführt. Ein häufig verwendetes Variationsprinzip ist das Galerkin Verfahren, das jedoch für bestimmte Gleichungstypen Nachteile aufweist. In dieser Arbeit wird die Eignung eines alternativen Variationsprinzips, der Least-Squares Finite Elemente Methode, zur Lösung von instationären Fluid-Struktur Interaktionsproblemen untersucht.
Dazu werden verschiedene Formulierungen für die Navier-Stokes Gleichungen, die in der Literatur vorgeschlagen wurden auf ihre Genauigkeit hin untersucht. Diese Formulierungen für die Navier-Stokes Gleichungen werden danach zunächst mit einer klassischen Galerkin Formulierung für die Struktur gekoppelt. Um eine reine Least-Squares Formlierung des gekoppelten Systems zu erhalten wird eine neue Formulierung für die instationren Gleichungen der linearen Elastizität entwickelt und analysiert. Hiermit wird dann eine stark gekoppelte reine Least-Squares Formulierung für Fluid-Struktur Probleme entworfen.
Die verschiedenen entwickelten Formulierungen werden schließlich mit verschiedenen Beispielen auf ihre Genauigkeit und Effizienz hin untersucht.

Abstract

Different numerical methods have been proposed for the solution of partial differential equations (PDE). Most of them are based on a variational principle which recasts the PDE into an equivalent integral equation. One of the most common principles is the Galerkin method, which has some specific disadvantages for some types of PDE. In this work an alternative variational principle, the least squares finite element method, will be tested with respect to its application for transient fluid-structure interaction problems.
The accurracy of different formulations which were proposed for the Navier-Stokes equations in literature will be tested. In a next step these formulations will be coupled with a standard Galerkin approach for the structure. After that a new formulation for the linear equations of elastodynamics is developed and analysed with respect to its stability and accuracy. With this formulation it is possible do develop a pure least squares formulation for the strongly coupled fluid-structure problem.
Finally the different formulations are tested with respect to their accuracy and efficiency.

Danksagung

Bedanken möchte ich mich vor allem bei Herrn Prof. H.G. Matthies für die Betreuung dieser Arbeit und dafür, daß er mir mit einer Stelle als wissenschaftlicher Mitarbeiter die Möglichkeit zur Anfertigung dieser Arbeit gegeben hat. Ganz besonders möchte ich mich auch bei Herrn Prof. G. Starke aus Hannover für die Übernahme des Koreferats und die Mitarbeit in der Promotionskommision bedanken. Ebenso gilt mein Dank Herrn Prof. U. Golze und Herrn Prof. M. Magnor für ihre Beteiligung an der Promotionskommision.
Ganz herzlich möchte ich mich auch bei meinen ehemaligen Kollegen A.K.M. Fahimuddin, S.Fischer, T.P. Fries, D. Fulger, C.Heimann, A.Keese, Markus Krosche, Martin Krosche, D.Liu, M.Meyer, R.Niekamp, C. Oldiges, T.Srisupattarawanit, J.Steindorf, J.Weimar und E.Zander für viele interessante Gespräche und die nette Arbeitsatmosphäre am Institut für wissenschaftliches Rechnen bedanken.
Schließlich gilt mein Dank meinen Eltern für die moralische und gelegentliche finanzielle Unterstützung, die mir sehr geholfen hat.

Contents

1 Introduction

After a short motivation why fluid structure interaction problems are important, a short overview of the current state of research in this field will be given. A discussion of the contents of this work follows.

1.1 Motivation

Fluid structure interaction (FSI) problems appear in nearly all fields of sciences and engineering. Examples include the blood flow in arteries, flutter phenomena in airplane wings and the wind building interaction in civil engineering. The collapse of the Tacoma narrows bridge in 1940 was an impressive example of the hazardous effects which can be caused by an underestimation of these phenomena. Thus windtunnel experiments were and still are obligatory for applications which might exhibit dangerous interaction phenomena.

As windtunnel tests require a high experimental effort they are quite expensive. In the last decades the price for computing power has decreased exponentially. Therefore the accuracy of numerical simulations has been increased while the costs have been reduced. But still no aircraft and no larger building is built without previous tests in the windtunnel. This implies that there is still a demand for better algorithms and numerical methods with increased accuracy.

Many of the commercial codes are based on numerical methods which in their basic form have been developed 20 or 30 years ago. As these algorithms were focused on finding solutions on the machines which were available at that time, they are really fast on current computers. But the numerical methods have evolved in that time as well and it could be possible that modern methods offer advantages over these established numerical schemes.

Thus the objective of this work is the exploratory evaluation of one of these alternatives. In the last years the interest in first order least squares finite element method (LSFEM) has continuously grown due to some promising properties of this method. It offers a sound mathematical foundation and hence is amenable to a rigorous analysis. Other advantages are emphasised differently by the authors who work in the field of least squares finite element methods.

The group of Prof. Manteuffel and Prof. McCormick, which has published several articles in this field, has set the focus on efficient multigrid algorithms (cf. [73]) for the solution of the resulting systems of equations (cf. [44], [102], [31], [95], [98]). In contrast Prof. Jiang sees the advantage of the least squares FEM mainly in its inherent robustness (cf. [83]). Both are one in demanding that the method should work with standard finite elements. This requirement can often be satisfied by using *augmented formulations* (cf. subsection 3.5.3). But other authors have also pointed out that these augmented formulations do have specific disadvantages (cf. [12]). Hence in other articles special elements are used to get stable numerical methods for unaugmented formulations (cf. [40]).

Which of these approaches will come out to be the right way remains an open question. But the published articles show a potential for new efficient numerical methods which motivate a detailed study of the least squares FEM for FSI problems.

1.2 Current State of Research

In this section an overview about the current state of research in the area of Fluid-Structure problems will be given. First the general ways to solve FSI problems will be explained with the focus on the special difficulties of FSI problems. After that the least squares methods will be analysed with respect to their possible use for solving FSI problems. The overview given here is focused on the essentials, but a more detailed review of the literature can be found in [86].

1.2.1 Fluid-Structure Interaction (FSI) Problems

A general FSI problem consists of a fluid part and a structural part which interact in some way. Usually the interaction is due to a common boundary between both parts, where some physical properties like momentum or energy can be exchanged. The discussion in this work will be limited to problems where the forces and deformations of the two parts are transferred along the boundary. This class of problems includes wind induced oscillations of buildings or bridges (cf. [79], [127], [107]) and simplified models of blood flow in elastic arteries (cf. [108], [76]).

Several difficulties are present in these FSI problems, which include the numerical solution of fluid problems, the numerical approximation of structural problems, the correct treatment of the coupling conditions and the grid deformation which is necessary for the simulation of fluid flows in deforming geometries.

For the solution of FSI problems *partitioned* and *monolithic* approaches can be differentiated. The partitioned approach uses specialised and hence often verified and effective solvers for the subproblems and couples these solvers in a suitable way to solve the coupled problem (cf. [103] for example). The partitioned approaches can be further subdivided into *staggered schemes* and *implicit schemes*.

Staggered schemes provide a convenient coupling method for transient problems. The timesteps of the subproblems are nested in a special manner and the loads and deformations are transferred between the subproblems after each step. In their basic form these schemes are usually only conditionally stable (cf. [106]). To overcome these difficulties, more sophisticated schemes were proposed (cf. [46]).

Alternatively the solvers for the subproblems could be embedded into special iteration schemes (cf. [100], [97], [128]). These procedures require the multiple solution of the subproblems in each timestep and thus imply higher computational costs for each timestep. But the higher computational costs for each timestep could be compensated by their advantages. In contrast to the staggered schemes they find the solution of the strongly coupled problem, and can be considered equivalent to the monolithic approaches (cf. [96]). Furthermore, the strong coupling enhances the stability properties and allows larger timesteps under certain conditions, which partially compensates the higher computational costs for each timestep.

Monolithic approaches aim at solving the nonlinear systems which describe the complete FSI problem. This approach offers several theoretical advantages (cf. [127], [80]). The nonlinearity can be efficiently treated by a standard Newton method. Additionally the fulfilment of conservation laws on the fluid structure interface can be usually ensured (cf. section 4.1 in [80]). Clearly this approach has some disadvantages as well. In each nonlinear iteration a linear system of equations has to be solved. As the different parts of this linear system of equations represent different materials, specialised preconditioners will be needed for an efficient solution of these systems. This would clearly make the solution process similar to the implicit schemes, which in an abstract framework could be considered as block preconditioners of the strongly coupled problem. Furthermore, the coupled code has to be implemented from scratch, and the introduction of new subproblems would always require a redesign.

1.2.2 Least Squares Methods

The chapters which discuss the use of the least squares method for the different subproblems are preceeded by an extensive review of the existing literature. Therefore this section will provide only a short overview about the existing literature. For a detailed introduction into the general ideas of the least squares FEM, the reader should consider the monographs [83] and [19].

As the numerical solution of the Stokes equations usually prepares the solution of the Navier-Stokes equations, several articles are available about the least squares FEM for these equations (cf. [49], [47], [15], [17], [21], [84], [58], [39], [45], [124], [22], [35] and more). An interesting detail is the fact that the Stokes equations for compressible fluids are equivalent to the equations of linear elasticity. Hence many articles treat these problems at the same time.

Nonetheless more specialised articles about the least squares FEM for the equations of linear elasticity have been published as well (cf. [25], [40], [38], [37], [135]). As the Galerkin FEM works quite well for the equations of linear elasticity, these articles are often focused on special problems like nearly incompressible solids. Interestingly the equations of linear elastodynamics have to our knowledge not been considered in any article yet. The same holds for nonlinear problems in elasticity.

Most of the formulations, which were developed for the Stokes equations, could be used for the Navier-Stokes equations as well. As the inclusion of convective terms usually does not lead to problems in the least squares FEM, the only requirement for a Stokes formulation to be usable for the Navier-Stokes equations is the availability of the velocity vector as unknown. Some articles about the Navier-Stokes equations are [8], [14], [20], [16], [21].

While at the beginning of this thesis, the field of least squares methods for FSI problems was completely empty, some publications appeared during its course. Probably the most important are several articles from the group around Profs. McCormick and Manteuffel (cf. [54], [53], [55], [76], [75], [77]). These are based on a series of papers [14],[20], [88], [44], [90], [43], and consider first order formulations for the Stokes, Navier-Stokes and linear elasticity equations which use the full displacement or velocity gradient as additional unknowns (cf. chapter 4, 5). These methods differ from the method proposed in this work by the used formulations. Furthermore most of the results in those articles are limited to stationary FSI problems. Only in [75] is the solution of a transient problem demonstrated.

Another short article which examines the solution of FSI problems with the least squares FEM is [91]. It proposes a partitioned approach in which the least squares FEM for the Navier-Stokes equations is coupled with a Galerkin structural formulation. In contrast to the results in this work the approach was partitioned and only stationary FSI problems were treated.

1.3 Structure of the Thesis

In the beginning the intention of this work was a bit different from what was finally achieved. The first idea was to implement a space time adaptive solution algorithm for FSI problems. This algorithm would have been based on a space time FE formulation for the strongly coupled problem. Inspired by the promising results presented in [83] the LSFEM was tried as a numerical method for this space time formulation. Although this approach seemed promising in the beginning and was proposed in a conference contribution [87], further tests for a larger publication did not succeed in confirming the expected convergence rates and revealed other problems. Hence the decision was to go back to a more established path of the LSFEM, namely the combination of a finite difference discretisation in time with the LSFEM for the spatial discretisation. The

examples in this work show that this path was not quite as established as originally thought.

The work starts with a brief review of the basic equations. That chapter mainly serves as a commented introduction of the notation which is used in continuum mechanics.

As the LSFEM still could be considered a non standard method, the fundamental ideas will be presented in chapter 3. Similarly to the Galerkin FEM, the essential concept will be the coercivity of the resulting bilinear forms. When continuity and coercivity can be shown in sufficiently strong norms, the LSFEM yields optimally accurate results. Unfortunately, the basic LSFEM does not always lead to coercive bilinear forms in the required norms. Several ways which have been proposed in the literature to overcome these problems will be discussed with their advantages and disadvantages. Additionally the treatment of nonlinear and transient problems will be illustrated. Both subproblems require a procedure which differs from that used in the standard Galerkin approach.

The following chapter is devoted to the numerical treatment of the fluid part with the least squares FEM. Although several articles have been published about this topic, still several questions remain. While the more mathematically oriented papers often do not go beyond the driven cavity as test case, the articles of the practitioners are sometimes not very rigorous in the verification. Hence this chapter intends to test the LSFEM against several benchmark problems. These include the flow around a cylinder (cf. [114]) and the Taylor vortex (cf. for example [72]). It comes out that various least squares formulations are capable of finding accurate solutions for these benchmark problems. Another issue which is still subject of controversial discussions is the lack of guaranteed global mass conservation. Two computational setups are used to analyse this issue. While the first problem which consists of a narrowing channel gives some impressions about the general behaviour, the second test revealed a strong dependence of the mass conservation on the shape of the domain. Corners have a significant effect on the overall quality of the solution. This might also explain the controversial results which have been presented in the literature (cf. [58], [22]).

In chapter 5 the solution of the equations of linear elastodynamics with the LSFEM is examined. Originally the solution of the structural problem with the LSFEM seemed to be straightforward. Computational tests quickly showed that a simple adaptation of the formulations which have been proposed for the equations of static linear elasticity did not work reliably. Thus the decision was to go back one step to the simpler wave equation which shows a similar behaviour. With this equation the basic concepts for the understanding of the LSFEM for transient problems are developed. An intermediate result shows that the stability of the least squares FEM for transient problems requires coercivity of the bilinear form for one timestep. By using this newly developed theorem two new stable least squares formulations for the equations of linear elastodynamics are developed. The first of these formulations augments the original problem by a seemingly redundant equation which ensures H^1 coercivity of the bilinear form. Therefore it could be used together with a usual finite element discretisation. The other formulation follows

a slightly different philosophy. Instead of augmenting the original equation to achieve H^1 coercivity, coercivity in a slightly weaker norm is accepted. For stability this makes the use of finite dimensional subspaces of $H(\mathrm{div})$ necessary. Both formulations are tested with several simple setups. The unaugmented formulation shows slightly better results.

To complete the coupled formulation, the two least squares functionals for the fluid and structural part are accompanied by two new functionals which represent the coupling conditions. This approach has also been proposed by the group around Prof. Manteuffel and Prof. McCormick (cf. [77]). But a rigorous analysis was not performed yet. Hence the chapter about the coupling conditions primarily aims at verifying the functional representing the coupled problem numerically. For this purpose different setups are developed which allow the separate evaluation of the two coupling conditions. Despite the missing theoretical foundation, the computational results indicate that the idea does indeed work as expected. Another new coupling scheme is proposed to implement a strong coupling between a Galerkin formulation for the structural part and a least squares formulation for the fluid part. For a simple test case both coupled formulations lead to similar results.

Finally chapter 7 brings the newly developed least squares methods closer to applications by showing results for a slightly more complex test case. Results which have been computed by other authors demonstrate that the test case shows a highly instationary behaviour. First the mixed Galerkin/least squares formulation is tested. In conjunction with high order elements for the fluid part, the formulation is capable of reproducing the results which have been found by other authors. The pure least squares formulation, which is tested afterwards, produces different results in dependence of the elements which are used for the discretisation of the structural part. For the lowest order structural elements, the strong numerical damping in the structure is the probable reason for the different results. Computations with the highest order element showed a quite different behaviour than the results from literature. A more thorough analysis is clearly required. Nonetheless these results demonstrate that the pure least squares formulation can cope with highly instationary problems.

A summary and an outlook on possible future directions of the LSFEM in the field of FSI problems close this work.

2 Mathematical Statement of Fluid Structure Problems

As the name already implies, fluid structure interaction (FSI) problems consist of some kind of fluid interacting with some type of structure. The basic fluid types can be differentiated into viscous or inviscid and compressible or incompressible behaviour. In some fields (like the simulation of blood flow) also other fluid types may appear. In this work the scope will be limited to viscous incompressible flows, which appear in a wide range of technical applications.

For the structural part most analyses focus on the mechanical or the thermal behaviour (for example in reentry vehicles). Here the focus will be on the mechanical behaviour. Although this narrows the area, there exists still an immense amount of possible choices for the structure. It could be modelled as plastic, elastic, compressible or nearly incompressible material. Here a model of a slightly compressible elastic material will be assumed in the structural part because it is a reasonable choice in the context of FSI problems.

Finally the coupling conditions must be considered. Fortunately, their basic form is the same for all choices of fluid and structural model. They have to ensure the compatibility of the velocities and the balance of the normal tractions along the interface.

To define the variables used, a short overview of the basic terms of continuum mechanics will be presented in the next sections. These equations will provide the basis for the fluid as well as for the structural part. A more detailed explanation can be found in ([74], [60], [56]).

2.1 Kinematics

Generally the domain of interest can be described as a set of points in a 3D Euclidian space. The set of points in the initial or undeformed state at a reference time $t = 0$ may be taken as the *material configuration*, which will be denoted by R_X and occupies the domain $\Omega(0)$. Any deformation can be described by a function φ, which maps an arbitrary point identified by its *material coordinate* $\boldsymbol{X}$ from the material configuration

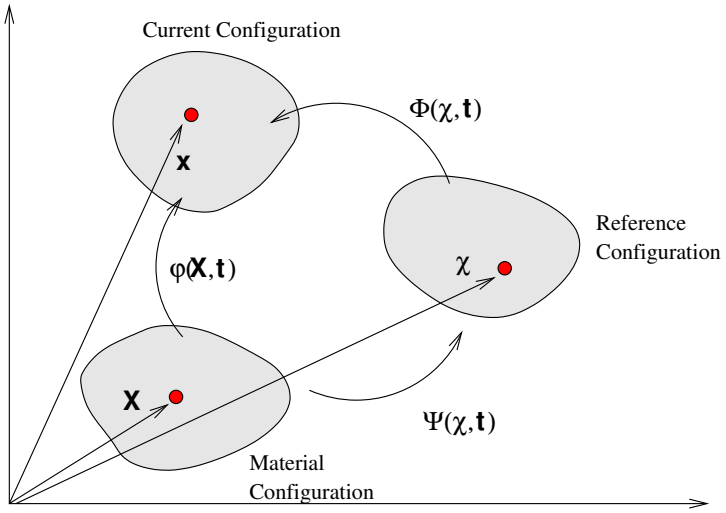

Figure 2.1. Material configuration, current configuration, reference configuration and the mappings between these configurations

to the *current configuration*, which will be named R_x (see also Fig. 2.1):

$$\varphi : R_X \times [0,T] \rightarrow R_x \times [0,T] \tag{2.1}$$
$$(\boldsymbol{X},t) \rightarrow \varphi(\boldsymbol{X},t) = (\boldsymbol{x},t). \tag{2.2}$$

Here $\boldsymbol{x}$ denotes the *spatial coordinate* and φ should be bijective to avoid self penetration. The domain $\Omega(t)$ is formed by the set of points at time t.

While the Lagrangian and Eulerian description use the material respectively the current configuration as starting point for the description of the kinematics, the arbitrary Lagrangian Eulerian (ALE) description introduces an arbitrary reference configuration R_χ, which is used as a basis for the description (cf. Fig. 2.1). A detailed discussion of the ALE formulation can be found in [60].

Between R_X and R_χ a mapping Ψ is introduced:

$$\Psi : R_X \times [0,T] \rightarrow R_\chi \times [0,T] \tag{2.3}$$
$$(\boldsymbol{X},t) \rightarrow \Psi(\boldsymbol{X},t) = (\boldsymbol{\chi},t). \tag{2.4}$$

A similar mapping, which maps the reference configuration into R_x will be called Φ:

$$\Phi : R_\chi \times [0,T] \rightarrow R_x \times [0,T] \tag{2.5}$$
$$(\boldsymbol{\chi},t) \rightarrow \Phi(\boldsymbol{\chi},t) = (\boldsymbol{x},t). \tag{2.6}$$

All mappings will be assumed to be bijective and continuous. The motion of a particle φ can then be expressed by $\varphi = \Phi \circ \Psi$ which shows that the mappings are not independent of each other.

The differences between the coordinates of a point in the three configurations will be denoted by:

$$\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{X}, \quad \boldsymbol{u}_\Phi = \boldsymbol{x} - \boldsymbol{\chi}, \quad \boldsymbol{u}_\Psi = \boldsymbol{\chi} - \boldsymbol{X} \tag{2.7}$$

with rates

$$\boldsymbol{v} := \dot{\boldsymbol{u}} = \left.\frac{\partial \boldsymbol{u}}{\partial t}\right|_X = \left.\frac{\partial \boldsymbol{x}}{\partial t}\right|_X = \dot{\varphi} \tag{2.8}$$

$$\boldsymbol{v}_\Phi := \dot{\boldsymbol{u}}_\Phi = \left.\frac{\partial \boldsymbol{u}_\Phi}{\partial t}\right|_\chi = \left.\frac{\partial \boldsymbol{x}}{\partial t}\right|_\chi = \dot{\Phi} \tag{2.9}$$

$$\boldsymbol{v}_\Psi := \dot{\boldsymbol{u}}_\Psi = \left.\frac{\partial \boldsymbol{u}_\Psi}{\partial t}\right|_X = \left.\frac{\partial \boldsymbol{\chi}}{\partial t}\right|_X = \dot{\Psi} \tag{2.10}$$

Depending on the configuration, three spatial differential operators will be used:

$$\begin{aligned} \nabla_\chi = \left(\frac{\partial}{\partial \chi_1}, \frac{\partial}{\partial \chi_2}, \frac{\partial}{\partial \chi_3}\right)^T, \quad \nabla_x = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)^T, \\ \nabla_X = \left(\frac{\partial}{\partial X_1}, \frac{\partial}{\partial X_2}, \frac{\partial}{\partial X_3}\right)^T. \end{aligned} \tag{2.11}$$

For the formulation of conservation laws on the reference domain, it is necessary to have an expression for the total time derivative in the reference and current configuration. Considering a scalar function $f^*(\boldsymbol{\chi}, t)$ which is defined for each point $\boldsymbol{\chi}$ in the reference configuration, there exist functions $f(\boldsymbol{x}, t)$ and $f^{**}(\boldsymbol{X}, t)$ which for each corresponding point $\boldsymbol{x}$ and $\boldsymbol{X}$ respectively have the same value in the other two configurations:

$$f(\boldsymbol{x}, t) = f^*(\boldsymbol{\chi}, t) = f^{**}(\boldsymbol{X}, t) \tag{2.12}$$

By using the chain rule the following equations can be found for the total time derivative in each of the three configurations (cf. [60]):

$$\frac{\partial f^{**}}{\partial t} = \frac{\partial f^*}{\partial t} + \boldsymbol{v}_\Psi \cdot \nabla_\chi f^* \tag{2.13}$$

$$\frac{\partial f^{**}}{\partial t} = \frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \nabla_x f \tag{2.14}$$

$$\frac{\partial f^{**}}{\partial t} = \frac{\partial f^{**}}{\partial t} \tag{2.15}$$

As ∇_χ has to evaluated in the reference domain, it is more convenient to express it in terms of the spatial gradient $\nabla_{\boldsymbol{x}}$. Using the fact that

$$\boldsymbol{v} - \boldsymbol{v}_\Phi = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\chi}} \boldsymbol{v}_\Psi \tag{2.16}$$

an alternative form of Eq. (2.13) is:

$$\frac{\partial f^{**}}{\partial t} = \frac{\partial f^{*}}{\partial t} + \nabla_x f \cdot (\boldsymbol{v} - \boldsymbol{v}_\Phi) \tag{2.17}$$

The vector $\boldsymbol{v}_\Phi$ denotes the velocity of the reference configuration in relation to the material configuration.

Within this framework the usual tensors can be defined. The *deformation gradient* $\mathbf{F}$ is defined as:

$$\nabla_X \varphi(X,t) = \mathbf{F} = \mathbf{I} + \mathbf{H}, \quad \mathbf{H} = \nabla_X \boldsymbol{u}. \tag{2.18}$$

Here $\mathbf{I}$ is the identity tensor.

The *material velocity gradient* is denoted by:

$$\nabla_X \boldsymbol{v}(\boldsymbol{X},t) = \dot{\mathbf{F}}(\boldsymbol{X},t). \tag{2.19}$$

and the *spatial velocity gradient* is:

$$\mathbf{L}(\boldsymbol{x},t) = \nabla_x \boldsymbol{v}(\boldsymbol{x},t). \tag{2.20}$$

2.1.1 Strain Measures

Strain measures should be invariant to rigid body motions and rotations as these are not of interest for the analysis of deformations. Hence the deformation gradient $\mathbf{F}$ is not well suited as a strain measure (cf. [56], [74]). If a strain measure satisfies these criteria, it is called *objective*.

Common objective strain tensors are:

- Right Cauchy-Green : $\mathbf{C} = \mathbf{F}^T\mathbf{F}$
- Green-Lagrange : $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$

For small deformations (i.e. $\mathbf{F} \approx \mathbf{I}$) the higher order terms can be neglected and lead to the linearised Green-Lagrange tensor ε:

$$\mathbf{E} = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T + \mathbf{H}^T\mathbf{H}) \approx \varepsilon = \frac{1}{2}(\mathbf{H} + \mathbf{H}^T) \tag{2.21}$$

The linearised Green-Lagrange tensor is not objective.

2.1.2 Strain Rates

The spatial velocity gradient $\mathbf{L}$ can be decomposed into a symmetric and an antisymmetric part (cf. Theorem 1.6 in [74]):

$$\mathbf{L} = \mathbf{D} + \mathbf{W} \tag{2.22}$$

with

$$\mathbf{D} = \frac{1}{2}(\mathbf{L} + \mathbf{L}^T) \tag{2.23}$$

$$\mathbf{W} = \frac{1}{2}(\mathbf{L} - \mathbf{L}^T). \tag{2.24}$$

Any rigid body motion is represented by $\mathbf{W}$ and makes $\mathbf{D} = 0$. Therefore the symmetric part $\mathbf{D}$ is well suited as a spatial strain rate tensor (cf. [74]).

A material strain rate tensor can be defined as follows:

$$\dot{\mathbf{E}}(\boldsymbol{X}, t) = \frac{1}{2}(\dot{\mathbf{F}}^T\mathbf{F} + \mathbf{F}^T\dot{\mathbf{F}}) \tag{2.25}$$

and is related to $\mathbf{D}$ by:

$$\dot{\mathbf{E}} = \mathbf{F}^T\mathbf{D}\mathbf{F}. \tag{2.26}$$

2.2 Conservation Laws

The equations which will be used for the description of FSI problems stem from basic physical principles which can be formulated as integral equations. Assuming sufficient regularity of the solution, the local formulations which are shown in the next sections can be derived.

Using the previously derived expressions for the total time derivative, the local forms of the conservation of mass and momentum become (cf. [60]):

$$\left.\frac{\partial \rho}{\partial t}\right|_{\chi} + (\boldsymbol{v} - \boldsymbol{v}_\Phi) \cdot \nabla_x \rho = -\rho(\nabla_x \cdot \boldsymbol{v}) \tag{2.27}$$

and

$$\rho\left(\left.\frac{\partial \boldsymbol{v}}{\partial t}\right|_{\chi} + ((\boldsymbol{v} - \boldsymbol{v}_\Phi) \cdot \nabla_x)\boldsymbol{v}\right) - \nabla_x \cdot \sigma = 0 \tag{2.28}$$

respectively. Here ρ is the mass density and σ the Cauchy stress tensor.

It can be shown that the conservation of rotational momentum implies the symmmetry of the stress tensor (cf. [74]):

$$\sigma = \sigma^T. \tag{2.29}$$

In the material description this reads:

$$\mathbf{F}\mathbf{P}^T = \mathbf{P}\mathbf{F}^T \tag{2.30}$$

where $\mathbf{P}$ is the 1st Piola-Kirchhoff stress tensor. Objective constitutive relations have to satisfy the balance of rotational momentum by ensuring the symmetry of the stress tensor.

A brief explanation and the definition of the stress tensors σ and $\mathbf{P}$ will follow in the next section.

2.3 Stress Measures

Using the balance of linear momentum for an imaginary tetrahedron with three surfaces parallel to the coordinate planes through a point x and the fourth surface with a unit normal vector $\boldsymbol{n}_C$, and letting the volume of this tetrahedron go to zero, it is possible to show a linear dependence between the normal vector $\boldsymbol{n}_C$ and the stress vector $\boldsymbol{t}_C$ of the current configuration (cf. [56]):

$$\boldsymbol{t}_C = \sigma \boldsymbol{n}_C. \tag{2.31}$$

The tensor σ in the equation above is called the *Cauchy* stress tensor, and represents the state of stress at a point. With the stress vector $\boldsymbol{t}_M$ in the material configuration, the 1st *Piola-Kirchhoff* stress tensor can be defined:

$$\boldsymbol{t}_M = \mathbf{P}\boldsymbol{n}_M. \tag{2.32}$$

The following relation between the 1st Piola-Kirchhoff and the Cauchy stress tensor can be established with the deformation gradient $\mathbf{F}$:

$$\mathbf{P} = (\det \mathbf{F})\sigma \mathbf{F}^{-T}. \tag{2.33}$$

For the formulation of objective material laws the symmetric *2nd Piola-Kirchhoff* stress tensor is better suited:

$$\mathbf{S} = \mathbf{F}^{-1}\mathbf{P} = (\det \mathbf{F})\mathbf{F}^{-1}\sigma \mathbf{F}^{-T}. \tag{2.34}$$

2.4 Constitutive Equations

The conservation of momentum is formulated in terms of the stress tensor, while deformations or velocities go into the strain tensor. This gap is closed by a *constitutive law*, which represents a model of the material behaviour.

2.4.1 Newtonian Fluid

Incompressible Newtonian fluids are described by a linear dependence between the Cauchy stress tensor and the symmetric part of the spatial velocity gradient:

$$\sigma = -p\mathbf{I} + 2\eta_f \mathbf{D} \tag{2.35}$$

with shear viscosity η_f and the pressure p.

2.4.2 Hookean Linear Elastic Solid

As already mentioned, the possible choices for modelling structures are enormous (cf. [137], [9] etc.). In this work only the most simple relation, which is called *Hookes law*, will be used:

$$\sigma = \lambda(\operatorname{tr}\varepsilon)\mathbf{I} + 2\mu_s\varepsilon \tag{2.36}$$

with the two Lamé constants λ and μ_s, which are related to Young's modulus E and the Poisson ratio ν_s by:

$$\lambda = \frac{\nu_s E}{(1+\nu_s)(1-2\nu_s)} \qquad \mu_s = \frac{E}{2(1+\nu_s)} \tag{2.37}$$

Although this constitutive relation is widely used, it is not objective (cf. [56]) and hence should not be used for large deformation analysis.

2.5 Coupled FSI Problem

In this section the resulting strong forms of the coupled FSI problem will be summarised.

2.5.1 Navier-Stokes Equations

Under the assumption of an incompressible fluid, the density ρ is constant in space and time. Hence the conservation of mass simplifies to $\nabla \cdot \boldsymbol{v} = 0$. Inserting Eq. (2.35) into Eq. (2.28) gives the Navier-Stokes equations for incompressible fluids:

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} + \rho(\boldsymbol{v} - \boldsymbol{v}_\Phi)\cdot\nabla_x\boldsymbol{v} - \eta_f\nabla_x\cdot(\nabla_x\boldsymbol{v}) + \nabla_x p = \boldsymbol{f} \quad \text{in} \quad \Omega \tag{2.38}$$

$$\nabla_x \cdot \boldsymbol{v} = 0 \quad \text{in} \quad \Omega. \tag{2.39}$$

The ALE formulation only influences the convective part through the prescribed grid velocities $\boldsymbol{v}_\Phi$.

These equations can be transformed into a normalised formulation, which is often more convenient for writing and analysis:

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} - \boldsymbol{v}_\Phi)\cdot\nabla_x\boldsymbol{v} - \nu_f\nabla_x\cdot(\nabla_x\boldsymbol{v}) + \nabla_x\tilde{p} = \tilde{\boldsymbol{f}} \quad \text{in} \quad \Omega \tag{2.40}$$

$$\nabla_x \cdot \boldsymbol{v} = 0 \quad \text{in} \quad \Omega. \tag{2.41}$$

In this form $\tilde{\boldsymbol{f}} = \boldsymbol{f}/\rho$, $\tilde{p}$ denotes the kinematic pressure and $\nu_f = \eta_f/\rho$ is used for the kinematic viscosity. Under certain conditions, the velocity field of the fluid may reach a stationary state. If this happens, the time derivative of the velocity field will vanish.

Hence the velocity field of this stationary state can be found with the Navier-Stokes equations for stationary flows of incompressible fluids (in the normalised form):

$$(\boldsymbol{v} \cdot \nabla_x \boldsymbol{v}) - \nu_f \nabla_x \cdot (\nabla_x \boldsymbol{v}) + \nabla \tilde{p} = \tilde{\boldsymbol{f}} \quad \text{in} \quad \Omega \tag{2.42}$$

$$\nabla_x \cdot \boldsymbol{v} = 0 \quad \text{in} \quad \Omega. \tag{2.43}$$

More details can be found in [74].

2.5.2 Equations of Linear Elastodynamics

Usually the structure is formulated in a Lagrangian description. Thus the convective parts in Eq. (2.28) become zero and inserting Eq. (2.36) into Eq. (2.28) yields:

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} - \mu_s \nabla_X \cdot (\nabla_X \boldsymbol{u}) - (\lambda + \mu_s) \nabla_X (\nabla_X \cdot \boldsymbol{u}) = \boldsymbol{f} \quad \text{in} \quad \Omega \tag{2.44}$$

with its stationary counterpart:

$$-\mu_s \nabla_X \cdot (\nabla_X \boldsymbol{u}) - (\lambda + \mu_s) \nabla_X (\nabla_X \cdot \boldsymbol{u}) = \boldsymbol{f} \quad \text{in} \quad \Omega. \tag{2.45}$$

These equations are commonly used as a basis for the development of least squares formulations for structural problems, which will be discussed in more detail in chapter 5.

2.5.3 Coupling Conditions

As it was shown in section 2.3 at each point of the domain a nominal stress vector $\boldsymbol{t}_M$ exists, which represents the forces acting along the axes of the coordinate system. Due to the balance of forces, this leads to the following coupling condition, which must be satisfied along the interface Γ_{if} between the fluid domain and the structural domain:

$$\sigma_f \boldsymbol{n}_C = \sigma_s \boldsymbol{n}_C \quad \text{on} \quad \Gamma_{if}(t). \tag{2.46}$$

Here $\boldsymbol{n}_C$ can be chosen either to be the outward normal vector of the fluid domain or the structural domain in the current configuration. Eq. (2.46) is often called the *traction* coupling condition.

Equivalently it could be formulated in the material configuration in terms of the 1st Piola Kirchhoff stress:

$$\mathbf{P}_f \boldsymbol{n}_M = \mathbf{P}_s \boldsymbol{n}_M \quad \text{on} \quad \Gamma_{if}(0). \tag{2.47}$$

where $\boldsymbol{n}_M$ denotes the outward normal vector in the material configuration. As the structure will be formulated in material coordinates, and the fluid in spatial coordinates, $\mathbf{P}_f$ must be expressed by σ_f. Using Eq. (2.33), Eq. (2.47) becomes:

$$(\det \mathbf{F}) \sigma_f \mathbf{F}^{-T} \boldsymbol{n}_M = \mathbf{P}_s \boldsymbol{n}_M \quad \text{on} \quad \Gamma_{if}(0). \tag{2.48}$$

If the fluid is viscous, the fluid sticks to the structure at the interface. Thus the velocities of the structure and the fluid must be the same at the interface:

$$\boldsymbol{v}_f = \dot{\boldsymbol{u}}_s \quad \text{on} \quad \Gamma_{if}(t). \tag{2.49}$$

The vector $\boldsymbol{u}_s$ denotes the displacements in the structural part. This condition is normally called the *compatibility* condition. For inviscid fluids, only the normal velocity has to be the same along the interface.

3 Least Squares Finite Element Methods

After a short introduction of the notation used, the basic idea of the least squares finite element method (LSFEM) and some simple error estimates for it will be presented. It will come out that optimal error estimates with respect to the used finite element functions require coercivity of the least squares functional with the norms in which the used finite elements achieve optimal interpolation results.

Several alternative methods, which follow the basic minimisation idea and hence can be seen as special LSFEM methods, were proposed in literature to achieve coercive bilinear forms in cases where the standard L^2 minimisation does not succeed. Section 3.5 provides a small overview about these methods.

Finally, a discussion of the current state of research regarding the LSFEM for initial boundary value problems is presented, followed by a small summary of the potential advantages of the LSFEM.

3.1 Notation

In the following sections the domain of the problems considered will be called $\Omega \subset \mathbb{R}^d$, with $d = 1,2,3$, and assumed to have a sufficiently smooth boundary Γ. If different boundary conditions have to be applied on different parts of the boundary, the parts of the boundary will be differentiated by a numerical or character index like Γ_n or Γ_1.

Throughout this work the following notation, which is based on the notation of [70], will be used. The space of square integrable functions defined on the domain Ω is denoted by $L^2(\Omega)$ and has the scalar product

$$(u, v)_{0,\Omega} := \int_\Omega u \cdot v \, d\Omega, \tag{3.1}$$

which induces the following norm

$$\|u\|_{0,\Omega} := (u, u)_{0,\Omega}^{1/2} = \left(\int_\Omega |u|^2 \, d\Omega \right)^{1/2}. \tag{3.2}$$

More generally the $L^p(\Omega)$ spaces are equipped with the norm

$$\|u\|_{0,p,\Omega} := \left(\int_\Omega |u|^p \, d\Omega \right)^{1/p} \tag{3.3}$$

Weak derivatives will be denoted by $D^{(\alpha)}$

$$D^{(\alpha)} f = \frac{\partial^{\alpha_1 + \cdots + \alpha_d} f}{\partial x_1^{\alpha_1} \cdots x_d^{\alpha_d}} \tag{3.4}$$

where $\alpha \in \mathbb{N}_0^d$ is a multiindex. The sum of its elements will be written $|\alpha| = \alpha_1 + \cdots + \alpha_d$. Then the definition of the *Sobolev* spaces $W^{k,p}(\Omega)$ is

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) \,|\, D^{(\alpha)} u \in L^p(\Omega) \quad \forall |\alpha| \leq k\}. \tag{3.5}$$

The norm on $W^{k,p}(\Omega)$ is defined by:

$$\|u\|_{k,p,\Omega} = \left(\sum_{|\alpha| \leq k} \|D^{(\alpha)} u\|_{0,p,\Omega}^p \right)^{1/p}. \tag{3.6}$$

The definition of spaces $W^{k,p}(\Omega)$ with $k \in \mathbb{R}$ is more difficult and non-unique and hence will be left out. In the places where these spaces appear, the definition given in [70] will be used.

Usually $H^k(\Omega) := W^{k,2}(\Omega)$. A seminorm on $H^k(\Omega)$ is defined by:

$$|u|_{k,\Omega} = \left(\sum_{|\alpha| = k} \|D^{(\alpha)} u\|_{0,2,\Omega}^2 \right)^{1/2}. \tag{3.7}$$

For some purposes also the space $H_0^k(\Omega)$ which is the completion of $C_0^\infty(\Omega)$ in the H^k norm is required.

In some parts of this work the following spaces will also be used:

$$H(\mathrm{div}; \Omega) := \{\boldsymbol{u} \in L^2(\Omega)^d \,|\, \nabla \cdot \boldsymbol{u} \in L^2\} \tag{3.8}$$

with the corresponding norm:

$$\|\boldsymbol{u}\|_{H(\mathrm{div};\Omega)} = (\|\boldsymbol{u}\|_{0,\Omega}^2 + \|\nabla \cdot \boldsymbol{u}\|_{0,\Omega}^2)^{1/2}. \tag{3.9}$$

For spaces $H_0^k(\Omega)$ the dual spaces are denoted by $H^{-k}(\Omega)$ and equipped with the following norm (the dual norm):

$$\|u\|_{-k,\Omega} = \sup_{0 \neq v \in H_0^k(\Omega)} \frac{\langle u, v \rangle_\Omega}{\|v\|_{k,\Omega}} \tag{3.10}$$

where $\langle \cdot, \cdot \rangle_\Omega$ denotes the duality pairing:

$$\langle u, v \rangle_\Omega := \int_\Omega u \cdot v \, d\Omega \tag{3.11}$$

Furthermore the following negative norms will be needed:

$$\|u\|_{-k^*,\Omega} = \sup_{0 \neq v \in H^k(\Omega)} \frac{\langle u, v \rangle_\Omega}{\|v\|_{k,\Omega}} \tag{3.12}$$

$$\|u\|_{-1,D,\Omega} = \sup_{0 \neq v \in H^1_D(\Omega)} \frac{\langle u, v \rangle_\Omega}{\|v\|_{1,\Omega}} \tag{3.13}$$

where $H^1_D(\Omega) \subset H^1(\Omega)$:

$$H^1_D(\Omega) = \{v \in H^1(\Omega) | v = 0 \quad \forall \boldsymbol{x} \in \Gamma_D\} \tag{3.14}$$

with Γ_D being a nonempty subset of the boundary of the domain Ω.

Let $\mathbf{V}(\Omega)$ denote a Hilbert space. Then $L^p(0, T; \mathbf{V}(\Omega))$ denotes the space of functions $u(t)$ with

$$\left(\int_0^T \|u(t)\|^p_{\mathbf{V},\Omega} \, dt \right)^{1/p} \leq \infty. \tag{3.15}$$

The corresponding norm is:

$$\|u\|_{L^2(0,T;\mathbf{V}(\Omega))} = \left(\int_0^T \|u(t)\|^p_{\mathbf{V},\Omega} \, dt \right)^{1/p}. \tag{3.16}$$

These spaces appear in section 6.3.

Furthermore the Gâteaux derivative $\mathcal{D}$ of functionals $\mathcal{F}(u)$ with respect to v is required:

$$\mathcal{D}\mathcal{F}(u)[v] = \lim_{\varepsilon \to 0} \frac{\mathcal{F}(u + \varepsilon v) - \mathcal{F}(u)}{\varepsilon} \tag{3.17}$$

The equations will be written completely in terms of the nabla-operator. This notation is unique except for the cross product, which leads to the *curl* operator and has two interpretations in two dimensions. If applied to a scalar (the z-component of some vector field with $u_1, u_2 = 0$), it should have the following meaning:

$$\nabla \times u = \begin{pmatrix} \frac{\partial u}{\partial y} \\ -\frac{\partial u}{\partial x} \end{pmatrix}. \tag{3.18}$$

Together with a vector $\boldsymbol{u} = (u_1, u_2)^T$ (from a vector field, where $u_3 = 0$) it is defined as follows:

$$\nabla \times \boldsymbol{u} = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \tag{3.19}$$

3.2 Abstract Framework

This section will show the basic idea of the least squares FEM. It is based on the monograph [83] and the article [19] which provide a good overview about the research activities in the field of least squares methods.

First a definition, which is one cornerstone of the theory of elliptic systems will be presented:

Definition 1 *Let $\mathbf{V}(\Omega)$ be a real Hilbert space with an associated norm $\|\cdot\|_{\mathbf{V},\Omega}$. A bilinear form $\mathcal{B}(u,v)$ is continuous if*

$$|\mathcal{B}(u,v)| \leq c_1\|u\|_{\mathbf{V},\Omega}\|v\|_{\mathbf{V},\Omega} \quad \forall u,v \in \mathbf{V}(\Omega)$$

It is said to be V-coercive if the following condition is satisfied:

$$|\mathcal{B}(u,u)| \geq c_2\|u\|^2_{\mathbf{V},\Omega} \quad \forall u \in \mathbf{V}(\Omega), c_2 > 0 \tag{3.20}$$

where c_1 and c_2 are independent of u and v.

Remark: The term *coercive* is used differently in literature. In [105], a bilinear form satisfying Eq. (3.20) and being continuous is called V-elliptic. This expression is used in [70] for bilinear forms satisfying only Eq. (3.20). Pedersen uses the term *coercive* in [105] for a definition which is named *weakly coercive* by other authors like in [109]. Throughout this work the term coercive will mean that a bilinear form satisfies Eq. (3.20). □

In the next paragraphs the following system of linear partial differential equation in a domain Ω with appropriate boundary conditions on Γ will be considered

$$\mathcal{L}\boldsymbol{u} = \boldsymbol{f} \quad \text{in} \quad \Omega \tag{3.21}$$
$$\mathcal{R}\boldsymbol{u} = \boldsymbol{g} \quad \text{on} \quad \Gamma \tag{3.22}$$

If there are Hilbert spaces $\mathbf{Y}(\Omega)\times\mathbf{Y}(\Gamma)$ and $\mathbf{V}(\Omega)$ such that $(\mathcal{L},\mathcal{R})$ is a homeomorphism (i.e. the mapping $\boldsymbol{u} \to (\mathcal{L}(\boldsymbol{u}),\mathcal{R}(\boldsymbol{u}))$ is a homeomorphism $\mathbf{V}(\Omega) \to \mathbf{Y}(\Omega)\times\mathbf{Y}(\Gamma)$) then the system Eqs. (3.21–3.22) is well posed in these Hilbert spaces and the a priori estimate

$$\|\boldsymbol{u}\|_{\mathbf{V},\Omega} \leq c_1(\|\mathcal{L}\boldsymbol{u}\|_{\mathbf{Y},\Omega} + \|\mathcal{R}\boldsymbol{u}\|_{\mathbf{Y},\Gamma}) \tag{3.23}$$

holds (cf. [19]).

The general idea of the LSFEM is to seek a function $\boldsymbol{u} \in \mathbf{V}(\Omega)$ which minimises a functional $\mathcal{J}(\boldsymbol{u})$ which consists of the square of the residuum $\mathcal{L}\boldsymbol{u}-\boldsymbol{f}$ and $\mathcal{R}\boldsymbol{u}-\boldsymbol{g}$ in the norms which are indicated by the a priori estimate Eq. (3.23):

$$\mathcal{J}(\boldsymbol{u}) = \frac{1}{2}\left(\alpha_1\|\mathcal{L}\boldsymbol{u}-\boldsymbol{f}\|^2_{\mathbf{Y},\Omega} + \alpha_2\|\mathcal{R}\boldsymbol{u}-\boldsymbol{g}\|^2_{\mathbf{Y},\Gamma}\right) \tag{3.24}$$

with some positive weights α_1 and α_2.

At the minimum of $\mathcal{J}$, the first variation must be zero:

$$\mathcal{D}\mathcal{J}(\boldsymbol{u})[\boldsymbol{v}] = 0 \quad \forall v \in \mathbf{V}(\Omega). \tag{3.25}$$

This is equivalent to seeking a $\boldsymbol{u} \in \mathbf{V}(\Omega)$ such that:

$$\mathcal{B}(\boldsymbol{u}, \boldsymbol{v}) = \mathcal{F}(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \mathbf{V}(\Omega) \tag{3.26}$$

with the bilinear forms:

$$\mathcal{B}(\boldsymbol{u}, \boldsymbol{v}) = \alpha_1(\mathcal{L}\boldsymbol{u}, \mathcal{L}\boldsymbol{v})_{\mathbf{Y},\Omega} + \alpha_2(\mathcal{R}\boldsymbol{u}, \mathcal{R}\boldsymbol{v})_{\mathbf{Y},\Gamma} \tag{3.27}$$

$$\mathcal{F}(\boldsymbol{v}) = \alpha_1(\boldsymbol{f}, \mathcal{L}\boldsymbol{v})_{\mathbf{Y},\Omega} + \alpha_2(\boldsymbol{g}, \mathcal{R}\boldsymbol{v})_{\mathbf{Y},\Gamma}. \tag{3.28}$$

Introducing finite dimensional subspaces $\mathbf{V}_h(\Omega)$ of $\mathbf{V}(\Omega)$ gives then a method suited for the numerical solution of partial differential equations. If the space $\mathbf{V}_h(\Omega)$ is not a subspace of $\mathbf{V}(\Omega)$, the method is called *nonconforming*. Some proposed least squares formulations use this approach because conforming finite elements would have numerical disadvantages (cf. [19], [21]).

The boundary conditions can be treated in two different ways (cf. [117]). Either the functional (3.24) is used directly for minimisation, or the function space $\mathbf{V}(\Omega)$, in which the solution is sought, is replaced by:

$$\tilde{\mathbf{V}}(\Omega) = \{\boldsymbol{u} \in \mathbf{V}(\Omega) | \mathcal{R}\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma\}. \tag{3.29}$$

Functions in $\tilde{\mathbf{V}}(\Omega)$ already satisfy the desired boundary conditions and hence the residuum of Eq. (3.22) will automatically be zero. Advantages and disadvantages of these two approaches will be discussed later.

To obtain a method which can be used for the numerical solution of partial differential equations, two conditions have to be satisfied referring to [19]:

- Optimality
- Practicality

Optimality means that there is a constant c_2 independent of h such that:

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{V},\Omega} \leq c_2 \inf_{v_h \in \mathbf{V}_h(\Omega)} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\mathbf{V},\Omega} \tag{3.30}$$

For conforming methods the optimality can be deduced from the a priori estimate Eq. (3.23). As $\mathbf{V}_h(\Omega) \subset \mathbf{V}(\Omega)$, the a priori estimate also holds for the finite dimensional subspaces. Hence the bilinear form $\mathcal{B}$ will be coercive on the finite dimensional subspaces which implies existence and uniqueness of the solution (cf. theorem 11.5 in [105]). With arguments from the standard finite element theory the estimate Eq. (3.30) can then be shown.

In [19] Bochev defines practicability by demanding that the inner products in $\mathcal{B}$ are computable, that standard $C^0(\Omega)$ finite elements can be used and that the resulting algebraic system of equations has a manageable condition number.

Standard $C^0(\Omega)$ finite elements consist of piecewise polynomials of a certain degree on some basic geometric domains (triangles, quadrilaterals, etc.) which form a partition of the domain Ω. These finite dimensional spaces are constructed in a way that the functions in that space are continuous. Continuity of the functions is sufficient to get a finite dimensional subspace of $H^1(\Omega)$ (cf. p.99 in [30]). It implies that the derivatives of functions in these spaces are at least in $L^2(\Omega)$ according to Eq. (3.5).

Higher derivatives are not in L^2 and thus would require either a weaker norm or elements with higher continuity, like finite dimensional subspaces of $H^2(\Omega)$, which provide a continuous first derivative (cf. p. 103 in [30]). Finite dimensional subspaces of $H^2(\Omega)$ are usually inconvenient to implement for higher space dimensions and also have the disadvantage of leading to a worse condition number in the resulting discrete system of equations. A higher condition number reduces the convergence rates of iterative solution procedures and leads to more numerical effort.

Therefore, currently most research is focused on first order methods which minimise the L^2 norm of a first order differential operator applied to the solution. Methods which use other norms, or allow higher order differential operators can normally not be used for practical applications (except the methods introduced in section 3.5).

Some other finite element spaces, like the Raviart-Thomas spaces, have been used in conjunction with the LSFEM as well [40]. They are subspaces of the slightly more uncommon spaces like $H(\mathrm{div};\Omega)$, where, besides the function itself, only the divergence lies in L^2.

Due to the local support of normal finite elements, the discrete system can be assembled element by element (similar to the Galerkin FEM), which leads to a sparse system of equations. It should be noted that the resulting system of linear equations is always symmetric positive definite when the bilinear form $\mathcal{B}$ is coercive. This is due to the underlying minimisation problem.

3.3 Approximation Properties of Finite Elements

One part of the theoretical framework of the LSFEM are estimates which guarantee that the finite subspace $\mathbf{V}_h(\Omega)$ of the function space considered can approximate an element of the original space with arbitrary accuracy in some norm if the triangulation is made fine enough. Furthermore, these estimates give some information about the norm of the error, which normally depends on the typical element size of the triangulation and the polynomial degree of the shape functions on the element.

First a few definitions from [70] will be given. An d-simplex of $\mathbb{R}^d$ is the convex hull κ of $d+1$ points, called the vertices of κ. For a non-degenerate simplex, these are not

all located on a single hyperplane. Size and shape of an d-simplex can be controlled by

$$h_\kappa = \text{diameter of } \kappa \tag{3.31}$$

and

$$\rho_\kappa = \sup\{\text{diameter of } B; B \text{ is a ball contained in } \kappa\}. \tag{3.32}$$

The regularity of κ is measured by

$$\sigma_\kappa = h_\kappa/\rho_\kappa \tag{3.33}$$

For each $h > 0$, $\mathcal{T}_h$ denotes a triangulation of the domain Ω with a polyhedral boundary made of closed simplices κ with diameters bounded by h:

$$\Omega = \bigcup_{\kappa \in \mathcal{T}_h} \kappa, \quad h_\kappa \leq h \tag{3.34}$$

In the error estimates, the following definition from [70] is of importance:

Definition 2 *A family of triangulations $\mathcal{T}_h$ is said to be regular as h tends to zero if there exists a number $\sigma > 0$, independent of h and κ, such that:*

$$\sigma_\kappa \leq \sigma \quad \forall \kappa \in \mathcal{T}_h \tag{3.35}$$

In addition, $\mathcal{T}_h$ is said to be uniformly regular (or quasi-uniform) as h tends to zero, if there exists another constant τ such that:

$$\tau h \leq h_\kappa \leq \sigma \rho_\kappa \quad \forall \kappa \in \mathcal{T}_h \tag{3.36}$$

Now for a fixed $k \geq 1$ the standard finite element spaces are ([70]):

$$\mathbf{S}_{h,k}(\Omega) = \{v_h \in C^0(\Omega); v_h|_\kappa \in \mathbf{P}_k \quad \forall \kappa \in \mathcal{T}_h\} \tag{3.37}$$

$$\mathbf{S}^0_{h,k}(\Omega) = \mathbf{S}_{h,k}(\Omega) \cap H^1_0(\Omega) \tag{3.38}$$

where $\mathbf{P}_k$ denotes the space of polynomials up to order k. An approximation result for these spaces can be found in [70] for example:

Lemma 1 *Let $\mathcal{T}_h$ be a regular triangulation of Ω. Then there exists a $v_h \in \mathbf{S}_{h,k}(\Omega)$ which satisfies the following error estimate for a real numbers $p > 1$, s and integers m with $0 \leq m \leq s+1, 1 \leq s \leq k$*

$$|v - v_h|_{m,p,\Omega} \leq c_1 h^{s+1-m} |v|_{s+1,p,\Omega} \quad \forall v \in W^{s+1,p}(\Omega) \tag{3.39}$$

The constant $c_1 > 0$ is independent of h and v.

As this estimate requires that the solution is at least in $W^{2,p}(\Omega)$, another estimate is required for solutions with less regularity. For this purpose the projection operator $\mathcal{P}_h : W^{1,p}(\Omega) \rightarrow \mathbf{S}_{h,k}(\Omega)$ will be needed:

$$\int_\Omega \nabla(\mathcal{P}_h v - v) \cdot \nabla v_h \, d\Omega = 0 \quad \forall v_h \in \mathbf{S}_{h,k}(\Omega) \tag{3.40}$$

$$\int_\Omega (\mathcal{P}_h v - v) \, d\Omega = 0 \tag{3.41}$$

Then the following theorem provides the required approximation results (taken from p.101 of [70]):

Theorem 1 *Assume that Ω is a convex polygon. Let $\mathcal{T}_h$ be a uniformly regular triangulation of Ω and let the real numbers s and p be such that $0 \leq s \leq k$ and $1 \leq p \leq \infty$. For $k \geq 2$ or for $k = 1$ and $2 \leq p < \infty$, there exists a constant c_1, independent of h, such that the projection $\mathcal{P}_h$ satisfies the error estimate:*

$$\|v - \mathcal{P}_h v\|_{0,p,\Omega} + h|v - \mathcal{P}_h v|_{1,p,\Omega} \leq c_1 h^{s+1} \|v\|_{s+1,p,\Omega} \forall v \in W^{s+1,p}(\Omega) \tag{3.42}$$

For the LSFEM especially the following two inverse estimates, which were taken from [70], are of great importance:

Lemma 2 *Let l and p be real numbers with $1 \leq l, p \leq \infty$. Under the assumption that the triangulation $\mathcal{T}_h$ is uniformly regular there exists a constant $c_1 > 0$ independent of h such that:*

$$|v|_{1,l,\Omega} \leq c_1 h^{-1+\min(0,d/l-d/p)} \|v\|_{0,p,\kappa} \quad \forall v \in \mathbf{S}_{h,k}(\Omega). \tag{3.43}$$

Furthermore if m is a non-negative integer and $r \leq p$ or if the triangulation $\mathcal{T}_h$ is uniformly regular and $l > p$, there exists a constant c_2, independent of h, such that:

$$|v|_{m,l,\Omega} \leq c_2 h^{\min(0,d/l-d/p)} |v|_{m,p,\Omega} \quad \forall v \in \mathbf{S}_{h,k}(\Omega). \tag{3.44}$$

Here d denotes the spatial dimension. Together with Lemma 1 this justifies the replacement of stronger norms through the weaker L^2 norm in the finite dimensional case (cf. subsection 3.5.1).

3.4 Introductory Example

To illustrate the use of the abstract framework, a short example which is taken from [83] will be provided in this section to demonstrate the basic techniques. The problem will be the Poisson equation in 1D on a domain $\Omega = [0 \ldots 1]$:

$$-\frac{\partial^2 u}{\partial x^2} = f \quad \text{in} \quad \Omega \tag{3.45}$$

$$u(x,0) = u_0 \tag{3.46}$$

$$u(0,t) = u(1,t) = 0 \tag{3.47}$$

Introducing the flux $\partial u/\partial x$ as new unknown p, an equivalent first order system is:

$$-\frac{\partial p}{\partial x} = f \quad \text{in} \quad \Omega \tag{3.48}$$

$$p - \frac{\partial u}{\partial x} = 0 \quad \text{in} \quad \Omega \tag{3.49}$$

A standard L^2 least squares functional would then be:

$$\mathcal{J}(u,p) = \frac{1}{2}\left(\left\|-\frac{\partial p}{\partial x} - f\right\|_{0,\Omega}^2 + \left\|p - \frac{\partial u}{\partial x}\right\|_{0,\Omega}^2\right) \tag{3.50}$$

The corresponding bilinear form can be obtained without difficulties and reads:

$$\mathcal{B}((u,p),(v,q)) = \left(-\frac{\partial p}{\partial x}, -\frac{\partial q}{\partial x}\right)_{0,\Omega} + \left(p - \frac{\partial u}{\partial x}, q - \frac{\partial v}{\partial x}\right)_{0,\Omega} \tag{3.51}$$

with right hand side part:

$$\mathcal{F}(f,(v,q)) = \left(f, -\frac{\partial q}{\partial x}\right)_{0,\Omega} \tag{3.52}$$

Thus the variational problem is to find a $(u,p) \in \mathbf{V}(\Omega)$ such that

$$\mathcal{B}((u,p),(v,q)) = \mathcal{F}(f,(v,q)) \quad \forall (v,q) \in \mathbf{V}(\Omega) \tag{3.53}$$

Referring to the notation of section 3.2, the spaces $\mathbf{Y}(\Omega)$ and $\mathbf{Y}(\Gamma)$ will be $L^2(\Omega)^2$ and $H^{1/2}(\Gamma)^2$ respectively, while the solution space $\mathbf{V}(\Omega)$ has to be determined. If an a priori estimate Eq. (3.23) is not available, the $\mathbf{V}$-coercivity of the bilinear form $\mathcal{B}$ has to be shown.

Theorem 2 *For all $(u,p) \in H_0^1(\Omega) \times H^1(\Omega)$ exists a positive constant c_1 independent of (u,p) such that:*

$$\|u\|_{1,\Omega}^2 + \|p\|_{1,\Omega}^2 \leq c_1 \mathcal{B}((u,p),(u,p))$$

Furthermore there exists a positive constant c_2 independent of (u,p) such that:

$$\mathcal{B}((u,p),(v,q)) \leq c_2(\|u\|_{1,\Omega}\|v\|_{1,\Omega} + \|p\|_{1,\Omega}\|q\|_{1,\Omega}) \quad \forall (u,p),(v,q) \in H_0^1(\Omega) \times H^1(\Omega)$$

Proof: Can be found on p. 39 of [83] and is provided in section A.1 □

For a numerical method, a finite dimensional subspace $\mathbf{V}_h(\Omega) \subset H_0^1(\Omega) \times H^1(\Omega)$ has to be selected. The finite dimensional variational problem is then to find a $(u_h, p_h) \in \mathbf{V}_h(\Omega)$ such that:

$$\mathcal{B}((u_h,p_h),(v_h,q_h)) = \mathcal{F}(f,(v_h,q_h)) \quad \forall (v_h,q_h) \in \mathbf{V}_h(\Omega) \tag{3.54}$$

As $\mathbf{V}_h(\Omega)$ is a subspace of $\mathbf{V}(\Omega)$, theorem 2 together with the Lax-Milgram lemma implies the existence of a unique finite dimensional solution. This finite dimensional solution satisfies the orthogonality condition:

$$\mathcal{B}((u-u_h, p-p_h), (v_h, q_h)) = 0 \quad \forall (v_h, q_h) \in \mathbf{V}_h(\Omega) \tag{3.55}$$

The space $\mathbf{S}^0_{h,k}(\Omega) \times \mathbf{S}_{h,k}(\Omega)$ which is constructed from the usual finite element spaces $\mathbf{S}_{h,k}(\Omega)$ is one possible choice for $\mathbf{V}_h$. For this selection the following theorem, which is an adaption of theorem 12 in [21], provides an error estimate:

Theorem 3 *Let $(u,p) \in H^1_0(\Omega) \times H^1(\Omega)$ and $(u_h, p_h) \in \mathbf{S}^0_{h,k}(\Omega) \times \mathbf{S}_{h,k}(\Omega)$ be the solutions of Eq. (3.53) and Eq. (3.54) respectively and assume that $u \in H^{s+1}_0(\Omega)$, $p \in H^{s+1}(\Omega)$ and $q = \min\{k, s\}$. Then*

$$\|u-u_h\|_{1,\Omega} + \|p-p_h\|_{1,\Omega} \leq c_1 h^q (\|u\|_{q+1,\Omega} + \|p\|_{q+1,\Omega}) \tag{3.56}$$

Proof: Let $\boldsymbol{u} = (u,p)^T$ and $\boldsymbol{v} = (v,q)^T$ with their finite dimensional counterparts $\boldsymbol{u}_h$ and $\boldsymbol{v}_h$. Using Eq. (3.55) and theorem 2:

$$\begin{aligned} c_2 \|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{1,\Omega} &\leq \mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{u}-\boldsymbol{u}_h) && (3.57)\\ &= \mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{u}-\boldsymbol{v}_h) + \mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{v}_h-\boldsymbol{u}_h) && (3.58)\\ &= \mathcal{B}(\boldsymbol{u}-\boldsymbol{u}_h, \boldsymbol{u}-\boldsymbol{v}_h) \leq c_3 \|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega} \|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,\Omega} && (3.59) \end{aligned}$$

Since $\boldsymbol{v}_h$ is an arbitrary element of $\mathbf{S}^0_{h,k} \times \mathbf{S}_{h,k}(\Omega)$, it follows that:

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega} \leq c_4 \inf_{v_h \in \mathbf{V}_h} \|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,\Omega} \tag{3.60}$$

Now theorem 1 and $q = \min\{k,s\}$ imply the existence of $\boldsymbol{v}_h \in \mathbf{S}^0_{h,k}(\Omega) \times \mathbf{S}_{h,k}(\Omega)$ such that

$$\begin{aligned} \|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,\Omega} &= (\|u-v_h\|^2_{1,\Omega} + \|p-q_h\|^2_{1,\Omega})^{(1/2)} && (3.61)\\ &\leq c_5 h^q (\|u\|_{q+1,\Omega} + \|p\|_{q+1,\Omega}) && (3.62) \end{aligned}$$

□

This theorem shows optimal convergence rates for the least squares method applied to the diffusion equation in 1D. For higher order elements, the solution has to be sufficiently smooth to guarantee optimal error estimates. The crucial ingredient for the error analysis was the coercivity of the bilinear form in the correct norm. What the correct norm is depends on the used finite dimensional subspaces. For Raviart-Thomas elements (cf. [28]), which are a subspace of $H(\mathrm{div};\Omega)$, coercivity in this slightly weaker norm is sufficient for optimality. Interestingly already in 2D full H^1 coercivity cannot be shown any more for the standard least squares approach applied to the diffusion problem (cf. subsection 3.5.3).

3.5 Coercivity and Stabilisation

As was shown in the previous section, the coercivity in appropriate norms is necessary for optimal convergence rates. To establish these coercivity results involves two closely connected subproblems. One is to prove the required estimates mathematically and the other is to find a suitable functional for this purpose.

For the first subproblem two ways are regularly used in literature. The first one is limited to elliptic systems of Agmon-Douglis-Nirenberg (ADN) type (cf. [1]) and was proposed generally for the LSFEM by Kellog and Aziz in [3]. In this approach the ADN theory is utilised to establish an a priori estimate Eq. (3.23), which then can be used to get the desired error estimates. The ADN-Theory requires that the *principal part* of the equations which must be uniformly elliptic is accompanied by boundary conditions which satisfy the so called *complementing condition* (cf. [1]). If these conditions are satisfied for an operator $\mathcal{L}$, the ADN theory provides a set of a priori regularity estimates which relate the regularity of the right hand side to the regularity of the solution. These estimates are similar to the following simplified estimate, which should serve to illustrate the basic idea:

$$\|\boldsymbol{u}\|_{q,\Omega} \leq c_1 \|\boldsymbol{f}\|_{q-1,\Omega} \tag{3.63}$$

Here q is an indexing which denotes the regularity of the right hand side vector $\boldsymbol{f}$. Using the fact that $\mathcal{L}\boldsymbol{u} = \boldsymbol{f}$, the squared estimate becomes:

$$\|\boldsymbol{u}\|^2_{q,\Omega} \leq c_1^2 \|\mathcal{L}\boldsymbol{u}\|^2_{q-1,\Omega}. \tag{3.64}$$

which already is an estimate like Eq. (3.23). Thus the coercivity of the least squares functional is directly available. Setting $q = 1$ gives the standard least squares principle. Once the coercivity is established, the error analysis follows mainly the outline of section 3.4. Otherwise, the estimates have to be derived "by hand" using theorems from functional analysis.

Methods for the second subproblem can be interpreted differently and are accordingly also denoted differently in the literature. The goal is always to establish a coercivity estimate for the bilinear form coming from the proposed functional, which in conjunction with the Lax-Milgram lemma guarantees the well-posedness of the variational problem. This process can be seen as a *stabilisation procedure* (cf. section 9.4 and 8.3 in [109]) in compliance with the terms used in the normal Galerkin FEM. Similar to the normal FEM, some problems do not require this stabilisation and allow the straightforward application of the LSFEM principle. But for other problems several ways have been proposed in the literature:

- Weighted LSFEM
- H^{-1} LSFEM

- Augmented LSFEM
- Streamline Diffusion
- First Order Systems $\mathcal{LL}^*$ (FOSLL*) (cf. [36])

The first three will be discussed in the following sections, while details about the latter two variants of the LSFEM can be found in [89] and [36] respectively.

3.5.1 Weighted LSFEM

As already mentioned, for systems of partial differential equations, the LSFEM minimises the sum of the squared norms of the residuals of the different equations. This process can be seen as compromise solution between the different equations. Considering the Navier-Stokes equations as an example, the LSFEM will violate the conservation of mass a bit and the conservation of momentum a little bit as well. Hence weighting the parts of the equation with factors can emphasise different properties depending on the objectives.

Weighting appears in literature in several forms with different justifications:

- To replace a different norm (cf. [21], [19])
- Matrix weighting (cf. [113])
- Physical arguments (cf. [77], [8], [63],[130], [7])
- Local error reduction (cf. [82], [83])

Only the weighting to "imitate" a different norm comes from stabilisation, but the other ideas related to weighting will be presented here as well.

For some problems coercivity cannot be shown in adequate norms if the functional consists only of L^2 norms. A popular example is the $up\omega$-formulation of the Stokes equations. With the ADN theory the a priori estimate Eq. (4.9) was found. Setting the regularity index q in Eq. (4.9) to zero yields:

$$\begin{aligned}\|\omega\|_{1,\Omega} + \|p\|_{1,\Omega} + \|\boldsymbol{v}\|_{2,\Omega} \leq \\ c_1(\|\nu\nabla\times\omega + \nabla p\|_{0,\Omega} + \|\nabla\times\boldsymbol{v} - \omega\|_{1,\Omega} + \|\nabla\cdot\boldsymbol{v}\|_{1,\Omega})\end{aligned} \tag{3.65}$$

Thus H^1 norms would be a part of the corresponding least squares functional. Although the evaluation of these norms is theoretically possible, it would again require C^1 ansatz-spaces with the already mentioned disadvantages.

Here weights offer a way to circumvent this difficulty. The essential argument is the fact that all norms on discrete subspaces, which are the ones actually used in computations, are equivalent up to a constant but which depends on h or on the dimension of the subspace. By connecting estimate (3.39) with (3.43), it can be shown that the H^1 norm

and the L^2 norm of a discrete standard finite element function differ by a factor of h which characterises the element size:

$$c_1 h \|u_h\|_{0,\Omega} \leq \|u_h\|_{1,\Omega} \leq c_2 h \|u_h\|_{0,\Omega} \quad \forall u_h \in \mathbf{S}_{h,k} \tag{3.66}$$

Here c_1 and c_2 are constants, which are independent of the function u. Thus the H^1 norm can often be replaced by a weighted L^2 norm. For other norms like $\|\cdot\|_{1/2,\Gamma}$ similar equivalence relations can be found (cf. [26], [117]). One drawback of these weighted methods is that the weighting influences the condition number and often no efficient solution method is known [19]. It should also be noted that this was just a simplified presentation of the general idea. A real analysis would be more difficult as these weighted methods are in fact nonconforming (cf. [21]).

In [113] the weighting is not only applied to the different parts of the equation. Instead a complete symmetric positive definite weighting matrix is introduced, which allows a fine tuning of the LSFEM. The article considers a diffusion problem, which is transformed into the following first order system:

$$\mathcal{L}_1(p) := \frac{\partial p}{\partial x} \tag{3.67}$$

$$\mathcal{L}_2(u,p) := p + \frac{\partial u}{\partial x} \tag{3.68}$$

$$\tag{3.69}$$

The matrix weighted bilinear form, which is examined in [113], is then:

$$\mathcal{B}((u,p),(v,q)) = \frac{1}{2}\int_\Omega \begin{pmatrix} \mathcal{L}_1(q) \\ \mathcal{L}_2(v,q) \end{pmatrix}^T \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \mathcal{L}_1(p) \\ \mathcal{L}_2(u,p) \end{pmatrix} d\Omega \tag{3.70}$$

Clearly setting $\alpha_{11} = \alpha_{22} = 1$ would lead to the usual least squares formulation. Using modern symbolic mathematics software, optimal weighting parameters α_{ij} are derived from the analytical solution and it is shown that the correct weighting matrix can increase the accuracy in a test example significantly. But this technology seems to be in its infancies and due to its dependence on powerful symbolic mathematical software it might be impossible to find solutions for more complex equations. The situation is very similar to that encountered in stabilised methods, where the optimal choice of the stabilisation parameter is often not clear for more complex systems of equations (cf. [60]).

Another weighting is motivated by the physical quantities appearing in a system of coupled partial differential equations. As the residuals may have different physical units, these will also go into the residual norms. Thus normally the LSFEM functional will sum up different physical quantities, which could lead to a violation of the principle of scale invariance. Using a different length scale could change the numerical results, which is clearly an unwanted behaviour for a numerical method. To overcome these difficulties,

some authors introduced scales to bring all equation residuals to the same physical unit ([8], [63], [130], [7]).

A different intention for the use of weights can be found in the so called iteratively reweighted LSFEM, which was proposed by Jiang in [82]. Jiang considered pure convection problems in two dimensions with appropriate boundary conditions and a constant velocity field. The standard L^2-LSFEM gives slightly diffusive results for this type of problem. Using the argument that along the discontinuities of the solution the finite elements are not able to reproduce the solution anyway, Jiang reduces the weights on these elements in an iterative procedure.

3.5.2 H^{-1} LSFEM

Negative norms have mainly two purposes in the LSFEM, namely the reduction of the regularity demands on the solution and the right hand side, and bringing the right balance into the LSFEM functional to make the corresponding bilinear form coercive in the proper norms. For the Laplace operator the elliptic variational theory shows that a right hand side in H^{-1} is admissible. As the right hand side goes into the functional which is minimised, it has to be in L^2 for the normal LSFEM. Hence the regularity requirements are increased. Often this issue might be of minor importance, but in some applications like mechanics the right hand side might not satisfy this requirement. Using a negative norm instead leads then to reduced regularity demands on the right hand side and the solution.

From definition (3.10) it is easy to see that this norm cannot be evaluated directly, like the L^2 norm or the H^k norms for $k > 0$. Rachford et.al. were the first to use negative norms in a numerical method [111]. Later their idea was applied to the Navier-Stokes equations [71], transfered to the LSFEM [24] and since then has been used in several applications ([34], [12]). These methods are normally called after this norm H^{-1} *methods* or *negative norm methods*.

The recipe to evaluate the discrete negative norms is based on the Galerkin variational principle for the problem

$$-\Delta w + w = f \quad \text{in} \quad \Omega \tag{3.71}$$

$$w = 0 \quad \text{on} \quad \Gamma, \tag{3.72}$$

which in its weak form is equivalent to the solution of

$$(\nabla w, \nabla v)_{0,\Omega} + (w, v)_{0,\Omega} = \langle f, v\rangle_\Omega \quad \forall v \in H_0^1(\Omega), \tag{3.73}$$

where $f \in H^{-1}(\Omega)$ and $w, v \in H_0^1(\Omega)$. Denoting the solution operator by $\mathcal{S} : H^{-1}(\Omega) \to H_0^1(\Omega)$ with $\mathcal{S}f = w$, the following equation holds (cf. [24], lemma 2.1):

$$\|u\|_{-1,\Omega}^2 = \sup_{0 \neq v \in H_0^1(\Omega)} \frac{\langle u, v\rangle_\Omega^2}{\|v\|_{1,\Omega}^2} = \langle \mathcal{S}u, v\rangle_\Omega \tag{3.74}$$

By introducing a suitable finite dimensional subspace $\mathbf{V}_h(\Omega) \subset H_0^1(\Omega)$, a discrete solution operator $\mathcal{S}_h : H^{-1}(\Omega) \to \mathbf{V}_h(\Omega) \subset H_0^1(\Omega)$ can be defined, which is then used in the numerical method. Normally the discrete solution operator is replaced by a spectrally equivalent operator, which is easier to evaluate.

Some authors first define the discrete negative seminorm $|\cdot|_{-1,\Omega}$ which is then extended to a complete norm by adding a weighted L^2 norm (cf. [12], [45]). The weight consists of a parameter $\alpha > 0$ and a term h^2, which depends on the characteristic element size. Actually both approaches are equivalent, as the parameter α can be introduced in Eq. (3.71) as well. In the negative norm computations which will be presented in this work the approach based on Eq. (3.71) is preferred, as it avoids the difficulties with the mesh dependent weight.

Because the solution operator $\mathcal{S}_h$ is a full matrix in the discrete case, a direct application would be prohibitively expensive. Therefore, most authors propose iterative solvers together with appropriate preconditioners, which can lead to very efficient algorithms (cf. [12], [24]).

For some problems, which will be discussed in more detail in the subsection 4.2.1, the reduction of regularity assumptions on the solution is crucial.

3.5.3 Augmented LSFEM

If the bilinear form of the first order formulation stemming from the partial differential equation is not coercive at all or not coercive in the desired norm, adding seemingly redundant equations can sometimes restore coercivity.

Probably one of the simplest examples to demonstrate this effect is the LSFEM procedure for the standard Poisson equation. The following explanations are based on [83]. Considering the Poisson equation:

$$-\nabla \cdot \nabla u = f \quad \text{in} \quad \Omega \tag{3.75}$$

$$u = g \quad \text{on} \quad \Gamma \tag{3.76}$$

and assuming for simplicity homogeneous boundary conditions (i.e. $g = 0$), one possible equivalent first order formulation is

$$\boldsymbol{p} - \nabla u = 0 \quad \text{in} \quad \Omega \tag{3.77}$$

$$\nabla \cdot \boldsymbol{p} = -f \quad \text{in} \quad \Omega \tag{3.78}$$

$$u = 0 \quad \text{on} \quad \Gamma. \tag{3.79}$$

Thus the corresponding L^2 functional would be

$$\mathcal{J}(u, \boldsymbol{p}) = \frac{1}{2}(\|\nabla \cdot \boldsymbol{p} + f\|_{0,\Omega}^2 + \|\nabla u - \boldsymbol{p}\|_{0,\Omega}^2) \tag{3.80}$$

which is a mapping

$$\mathcal{J} : H(\mathrm{div};\Omega) \times H_0^1(\Omega) \to \mathbb{R}. \tag{3.81}$$

Let $\mathcal{B}((u,\boldsymbol{p}),(v,\boldsymbol{q}))$ denote the associated bilinear form. Then the following estimates hold [83]

$$c_1(\|u\|_{1,\Omega}^2 + \|\boldsymbol{p}\|_{H(\mathrm{div};\Omega)}^2) \leq \mathcal{B}((u,\boldsymbol{p}),(u,\boldsymbol{p})) \quad \forall (u,\boldsymbol{p}) \in H_0^1(\Omega) \times H(\mathrm{div};\Omega)^d \tag{3.82}$$

$$\mathcal{B}((u,\boldsymbol{p}),(u,\boldsymbol{p})) \leq c_2(\|u\|_{1,\Omega}^2 + \|\boldsymbol{p}\|_{H(\mathrm{div};\Omega)}^2) \quad \forall (u,\boldsymbol{p}) \in H_0^1(\Omega) \times H(\mathrm{div};\Omega)^d. \tag{3.83}$$

For standard finite element functions, this estimate implies suboptimal convergence rates because the standard finite elements do not have optimal approximation properties in $H(\mathrm{div};\Omega)$. Optimality with standard finite elements would require full H^1 coercivity of the bilinear form. Extending Eqs. (3.77– 3.79) by two additional equations:

$$\nabla \times \boldsymbol{p} = 0 \quad \text{in} \quad \Omega \tag{3.84}$$

$$\boldsymbol{n} \times \boldsymbol{p} = 0 \quad \text{on} \quad \Gamma \tag{3.85}$$

with $\boldsymbol{n}$ denoting the outward normal vector, results in a system of equations which is still equivalent to the system Eq. (3.75) and Eq. (3.76). Without going into detail, it is possible to show full H^1 coercivity for the pure L^2 functional which is based on the extended system (cf. [83]). This implies optimal convergence rates with standard finite elements.

But the augmented LSFEM is not as advantageous as it might seem. Actually the domain of the first order formulation Eqs. 3.77–3.79 is $H_0^1 \times H(\mathrm{div};\Omega)$. The additional equations lead to full H^1 coercivity but as a consequence the domain of the operator changes to $H_0^1(\Omega) \times H^1(\Omega)^d$, which is obviously smaller than the original domain. Hence, the augmented formulation finds only the projection of the solution to this smaller space which might be too small. Therefore the augmented formulations are not of great use as they further increase the regularity demands on the solution (cf. [18]). Similar problems occur also for some formulations of the Stokes and Navier-Stokes equations (cf. subsection 4.2.1).

3.6 LSFEM for Nonlinear Problems

Basically two ways can be considered to treat nonlinear partial differential equations. The usual approach, which is commonly used together with the Galerkin variational principle is to discretise the fully nonlinear equations. Considering an arbitrary nonlinear partial differential operator $\mathcal{L}$:

$$\mathcal{L}u = 0 \quad \text{in} \quad \Omega \tag{3.86}$$

the discretisation will yield a nonlinear algebraic system:

$$\mathbf{K}(\boldsymbol{u})\boldsymbol{u} = 0 \tag{3.87}$$

where the finite dimensional vector $\boldsymbol{u}$ represents the approximating finite element function. Usually this algebraic system is then solved with some appropriate method for nonlinear systems of equations.

The other approach is to linearise the differential operator $\mathcal{L}$ on the continuous level. In the following $\mathcal{L}$ should be a mapping from $\mathbf{V}(\Omega) \subset W^{k,p}(\Omega)^n$ to $\mathbf{X}(\Omega) \subset W^{k,p}(\Omega)^n$ and is assumed to be Fréchet differentiable. The exact solution $\boldsymbol{w}^* \in \mathbf{V}(\Omega)$ of the nonlinear problem satisfies (in a weak sense):

$$\mathcal{L}(\boldsymbol{w}^*) = 0 \tag{3.88}$$

In most cases it will not be possible to directly find a solution $\boldsymbol{w}^*$ of Eq. (3.88). Starting with an initial guess $\boldsymbol{w}_n$, Newton's method solves

$$D\mathcal{L}(\boldsymbol{w_n})[\delta \boldsymbol{w}_n] + \mathcal{L}(\boldsymbol{w}_n) = 0 \tag{3.89}$$

to find a new solution $\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \delta \boldsymbol{w}_n$. Clearly $\boldsymbol{w}_{n+1}$ will not be the exact solution, but it can be shown that under sufficiently strong assumptions on $\mathcal{L}$ and the initial guess $\boldsymbol{w}_0$ the sequence $\boldsymbol{w}_n$ converges to $\boldsymbol{w}^*$ (cf. [115]).

Assuming that Eq. (3.87) is solved with Newton's method, the system of linear equations which has to be solved in each iteration will be exactly the same as that found by discretising Eq. (3.89) with the Galerkin variational principle. Thus linearisation and discretisation could be exchanged in the Galerkin method. This observation does not hold for the least squares finite element method as the nonlinear terms will be squared in the latter.

To illustrate the differences between these approaches, the abstract variational statements will be derived by using both approaches. First the minimisation of the nonlinear functional coming from the L^2 LSFEM principle will be examined. The nonlinear least squares functional is:

$$\mathcal{J}(u) = \frac{1}{2}\|\mathcal{L}u\|_{0,\Omega}^2 \tag{3.90}$$

and its first variation:

$$\mathcal{D}\mathcal{J}(u)[v] \quad = \quad (\mathcal{D}\mathcal{L}(u)[v], \mathcal{L}(u))_{0,\Omega} \tag{3.91}$$

Now Newton's method has to be applied to $\mathcal{D}\mathcal{J}(u)[v]$ which is nonlinear in u. The first variation of $\mathcal{D}\mathcal{J}(u)[v]$ is:

$$\begin{aligned} \mathcal{D}^2\mathcal{J}(u)[v, \delta u] &= (\mathcal{D}^2\mathcal{L}(u)[v, \delta u], \mathcal{L}(u))_{0,\Omega} && (3.92)\\ &+ (\mathcal{D}\mathcal{L}(u)[v], \mathcal{D}\mathcal{L}(u)[\delta u])_{0,\Omega} && (3.93) \end{aligned}$$

Finally the integral statement for one Newton iteration reads:

$$\begin{aligned} (\mathcal{D}^2\mathcal{L}(u)[v, \delta u], \mathcal{L}(u))_{0,\Omega} + (\mathcal{D}\mathcal{L}(u)[v], \mathcal{D}\mathcal{L}(u)[\delta u])_{0,\Omega} + \\ (\mathcal{D}\mathcal{L}(u)[v], \mathcal{L}(u))_{0,\Omega} = 0 \quad \forall v \in \mathbf{V}(\Omega) \end{aligned} \tag{3.94}$$

Using the other approach, the LSFEM will minimise the functional which stems from the linearisation of Eq. (3.86) (i.e. Eq. (3.89)):

$$\mathcal{J}(\delta u) = \frac{1}{2}\|\mathcal{DL}(u)[\delta u] + \mathcal{L}(u)\|^2_{0,\Omega} \tag{3.95}$$

Its first variation leads then to the integral statement for one iteration:

$$(\mathcal{DL}(u)[\delta u], \mathcal{DL}(u)[v])_{0,\Omega} + (\mathcal{L}(u), \mathcal{DL}(u)[v])_{0,\Omega} = 0 \quad \forall v \in \mathbf{V}(\Omega). \tag{3.96}$$

Comparing Eq. (3.96) and Eq. (3.94) it is obvious that exchanging the order of linearisation and the application of the variational principle leads to different results. Eq. (3.94) possesses an additional term with the second Fréchet derivative.

Both ways were proposed and used in literature. Examples for the first approach can be found in [21], [12], [14] and [20]. In these works the analysis of some least squares formulations for the Navier-Stokes equations is performed on the fully nonlinear least squares functional. In contrast, Jiang (cf. [83]) proposes the second way. When the linearisation is the first step in the treatment of nonlinear problems, the least squares procedure is only applied to linear problems and the established theory can be applied. The only broader discussion of the differences between these methods which is known to the author, is the work [54]. Here both approaches are embedded into a multigrid solution algorithm. Referring to p.39 in [54] the term including the second Fréchet derivative is dominated by the lower order terms near the solution. This leads to the conclusion, that the radius of convergence will be smaller for the second approach.

If the initial guess is not within the ball of convergence around $\boldsymbol{w}^*$, the Newton algorithm does not converge at all. Especially for the Navier-Stokes equations it is well known that the ball of convergence becomes smaller with increasing Reynolds number. This will clearly reduce the probability that an arbitrary initial guess leads to a convergent solution. Therefore, in [83] Jiang proposes a continuation method for higher Reynolds numbers. The solution process starts with finding a solution for the problem at a relatively low Reynolds number. This result is then used as initial guess for the same problem at a slightly higher Reynolds number. Repeating this procedure, the Reynolds number can slowly be increased until the desired value is reached. Clearly this procedure is basically a homotopy method.

Another view could be based on the meaning of the iterates. While in the first approach the nonlinearity goes into the algebraic system of equations, the second approach iterates over approximations of functions. Therefore a sufficient approximation quality has to be ensured in each iteration to guarantee convergence. In this work the second approach will be preferred.

3.7 LSFEM for Initial Boundary Value Problems

For the normal Galerkin method which stems from a variational formulation, the separate treatment of the time derivative and the spatial derivatives is straightforward and is a property which is often essential for the convenient mathematical analysis.

A slightly different situation appears when using the LSFEM, because the differential operator and the time derivative of the weighting function influence the weighting term. Considering an arbitrary partial differential equations (PDE) of first order in space and time with initial and boundary conditions:

$$\begin{aligned} \frac{\partial u}{\partial t} + \mathcal{L}u &= f \quad \text{in} \quad \Omega, \qquad u(0) = u_0 && (3.97)\\ u &= 0 \quad \text{on} \quad \Gamma && (3.98) \end{aligned}$$

the most common approach in the LSFEM is to first discretise in time to get a spatially strong form (cf. [60]). After that, the LSFEM is applied to this strong form. Using a typical finite difference approximation for the time derivative, the class of θ-methods (cf. for example pp.162 in [126]) becomes for Eq. (3.97) in the homogeneous case ($f = 0$)

$$u_{n+1} + \Delta t\theta\mathcal{L}u_{n+1} = u_n - \Delta t(1-\theta)\mathcal{L}u_n \quad \text{in} \quad \Omega \tag{3.99}$$

The corresponding L^2 least squares functional is then:

$$\mathcal{J}(u_{n+1}, f_1) = \frac{1}{2}\|u_{n+1} + \Delta t\theta\mathcal{L}u_{n+1} - f_1\|^2_{0,\Omega} \tag{3.100}$$

with $f_1 = u_n - \Delta t(1-\theta)\mathcal{L}u_n$. A function $u_{n+1} \in \mathbf{V}(\Omega)$ which minimises the functional has to satisfy:

$$\begin{aligned} (u_{n+1} + \Delta t\theta\mathcal{L}u_{n+1}, v + \Delta t\theta\mathcal{L}v)_{0,\Omega} = \\ (u_n - \Delta t(1-\theta)\mathcal{L}u_n, v + \Delta t\theta\mathcal{L}v)_{0,\Omega}. \quad \forall v \in \mathbf{V}(\Omega) \end{aligned} \tag{3.101}$$

Now the question is how the variational formulation for the fully continuous formulation must look like. Clearly inserting the finite difference approximation for the time derivative should lead to Eq. (3.101). The only variational formulation which satisfies this criterion is:

$$\left(\frac{\partial u}{\partial t} + \mathcal{L}u, \frac{\partial v}{\partial t} + \mathcal{L}v\right)_{0,\Omega} = \left(f, \frac{\partial v}{\partial t} + \mathcal{L}v\right)_{0,\Omega} \qquad \forall v \in \mathbf{V}(\Omega). \tag{3.102}$$

Unfortunately, also the time derivative of the weighting function appears in the formulation, which complicates the analysis significantly. Therefore, the analysis of least squares formulations for initial boundary value problems has been done only for the resulting fully discrete equations so far. For the convective transport equation this was done in

[50] and [59]. An analysis of a time continuous least squares like method, where $\mathcal{L}v$ is used instead of $\frac{\partial v}{\partial t} + \mathcal{L}v$ can be found in chapter 16 of [125].

The family of θ-methods has been used successfully for the Navier-Stokes equations ([121], [120], [122]), convective transport ([33]), the incompressible Euler equations ([134]) and electromagnetic problems ([133]), to mention just a few applications. Thus there is some evidence that the stability properties of the θ-methods seem to be conserved if they are used in conjunction with the LSFEM. Thus choosing $\theta \geq 1/2$ should normally lead to a stable numerical scheme.

Higher order schemes in time were proposed mainly for purely convective problems (cf. [104]). The general idea is again to apply the time discretisation before using the LSFEM for the spatial discretisation (cf. [60]).

Another approach for solving initial boundary value problems with the LSFEM is the space-time LSFEM. This approach was even considered as a pure time integration scheme. A comparison of different time integration algorithms, which can be found in [137] and is based on works from the seventies [136], [132], shows that the space-time least squares approximation of transient problems yields a good accuracy.

For convection problems a space-time LSFEM was introduced by Nguyen and Reynen in 1984 ([101]). They also showed that this approach does not need any special treatment like upwinding or the Taylor-Galerkin method. But a later publication by Donea ([59]) showed that this method is less accurate and more dissipative than an approach based on using the θ-method for time integration.

In [8] the space-time LSFEM was used for the Navier-Stokes equations in the velocity-stress-pressure (cf. subsection 4.2.1 and subsection 4.2.2) formulation, and in [87] a similar scheme was proposed for the velocity-vorticity-pressure formulation (cf. subsection 4.2.1 and subsection 4.2.2). Unfortunately, in [8] the time stepping was only used to achieve the stationary state of a driven cavity example. Therefore the paper makes no statements about the time accuracy of the LSFEM.

A recent publication from Majidi and Starke ([93] and [94]) is about the space time LSFEM for parabolic problems. They want to use the space time LSFEM to utilise the "built-in" error indicator of the LSFEM. For the mathematical analysis they consider a numerical procedure which splits the space-time domain into time slabs. One of these slabs, consisting of only one element in time direction, is considered for the analysis. In the time direction linear shape-functions are inserted into the integral equations, which are then solved by Simpson's rule, which is exact for polynomials of order two and hence sufficient for the used shape functions.

In the first article [93] the stability of the resulting numerical method is analysed. The authors show coercivity of the bilinear form for one timestep and also stability for the complete instationary problem. Numerical tests and details about the adaptive algorithm are published in the second part [94]. The numerical results confirm the theoretical results.

Looking at the space-time LSFEM for the transport equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \quad \text{in} \quad \Omega \tag{3.103}$$

the following integral statement can be derived

$$\int_0^T \int_\Omega \left(c\left(\frac{\partial u}{\partial t}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial x}\frac{\partial v}{\partial t} \right) + \frac{\partial u}{\partial t}\frac{\partial v}{\partial t} + c^2 \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} \right) d\Omega\, dt. \tag{3.104}$$

Obviously the last two terms in Eq. (3.104) are similar to the terms obtained when applying the normal Galerkin method to an anisotropic diffusion equation. These terms can be interpreted as stabilisation in the streamline direction. But according to the results obtained in [59] the formulation is probably "over-stabilised". Normally, in numerics more stability is coupled with less accuracy (cf. [72]), which explains the low accuracy of the space-time LSFEM. Another disadvantage is the increased number of unknowns.

3.8 Advantages of the LSFEM

A good question concerning the LSFEM is why there is the need for another method, when the classical Bubnov-Galerkin scheme together with special element pairs (cf. [137], [72]) or stabilisation techniques (cf. [81], [109]) is also capable of solving the Navier-Stokes equations.

One argument for the LSFEM, which uses L^2 norms for minimisation, is the included stabilisation and the absence of the *Ladyzhenskaya-Babuška-Brezzi* (LBB) condition when it is used for mixed problems ([19], [83]). This condition regularly comes up within the Galerkin framework when saddle point problems have to be treated. Examples include mixed methods (cf. [30]), the Stokes equations (cf. [70]) and domain decompositions methods (cf. [110], [131]). Essentially the consequence of this condition is that the finite dimensional subspaces used for the discretisation of the mentioned problems have to satisfy a discrete LBB condition. These finite dimensional subspaces are difficult to find and sometimes inconvenient. As the LSFEM recasts the saddle point problems into equivalent minimisation problems, the LBB condition is circumvented.

This fact makes the LSFEM attractive for the solution of arbitrary PDEs because the LSFEM can be applied directly in a kind of black box manner. But as already explained in the previous sections, the price to pay for this uncomplicated behaviour is often suboptimal convergence.

Another advantage is that the resulting system of equations is always symmetric positive definite. Hence the use of more robust iterative solvers for symmetric problems, like the conjugate gradient method, is already effective when used in conjunction with some simple preconditioner like the Jacobi method (cf. [83]). Furthermore, if full H^1-coercivity can be established, even more efficient multigrid algorithms could be used for the solution of the resulting system of equations (cf. [31], [76], [43]).

Last but not least, the basic L^2 LSFEM provides a simple and very effective error indicator for adaptive methods which is simply the element residual measured in the L^2 norm. Several authors have shown for different problems that this indicator works very well. (cf. [10], [93], [94],[40]).

Whether these theoretical advantages give real benefits in engineering applications will be analysed in this work.

4 Numerical Treatment of the Fluid Domain

After a short discussion of possible methods to solve the incompressible Navier-Stokes equations, the focus is shifted to the possible least squares methods. Several methods, which were proposed in literature, are documented and then numerically evaluated with respect to their accuracy for several benchmark problems.

4.1 Numerical Procedures for the Navier-Stokes Equations

Neglecting the finite difference methods, all methods currently used for the numerical solution of the Navier-Stokes equations are based on a variational formulation. This variational formulation is obtained by multiplying the PDE with a suitable set of test functions and integrating over the domain (p.130 in [109]). The choice of the test functions determines the properties of the resulting method.

Using the characteristic function of small subdomains leads to the finite volume method (FVM), which historically evolved in the area of computational fluid dynamics. Inside one *control volume* the physical quantities of interest like mass or momentum are conserved (cf. section 2.5.4 in [66]), which motivates the term *local conservation*. This property is the main advantage of the FVM. But the FVM has also some disadvantages. The mathematical analysis of these methods is more difficult than that for the finite element methods, and it is difficult to construct higher order schemes (cf. section 2.6.2 [66]).

The discontinuous Galerkin methods (for an overview cf. [52]) keep the property of local flux conservation while having the same advantageous properties as the FEM. Therefore they can be easily extended to higher order. These features have led to a huge interest in these methods during the last years. A slight disadvantage of these methods is the higher number of unknowns.

In contrast to the FVM, which has its roots in the area of CFD, the finite element method (FEM) was originally proposed for the solution of structural problems (cf. [78],

[99], vol. 1 of [137]). Later the method was extended to other applications and also to the Navier-Stokes equations. Unfortunately, it is quite difficult to use the FEM for the Navier-Stokes equations, as is revealed by the mathematical analysis (cf. [70]). Convection terms and the saddle point structure of the Navier-Stokes equations are the main difficulties. While the latter problem can be addressed by either element-pairs satisfying the inf-sup condition or the use of stabilisation, the first difficulty always requires some kind of stabilisation (cf. chapter 8 in [109]). Methods to stabilise the FEM include the *streamline upwind Petrov Galerkin* (SUPG), *Galerkin least squares* (GLS) and other methods (cf. [109]). Advantages of the stabilised Galerkin FEM for the Navier-Stokes equations are the established mathematical theory, good accuracy and global conservation. Furthermore higher order methods can be implemented, although the right stabilisation parameters for these methods are still subject of active research (cf. p. 65, [60]). Disadvantages of the stabilised Galerkin FEM are the resulting nonsymmetric systems of equations, requiring iterative solvers for nonsymmetric systems of equations, which are less robust than those for symmetric problems.

Some advantages of the FEM, namely the well developed mathematical theory and the possibility to use elements of arbitrary order, transfer directly to the LSFEM. Compared to the stabilised Galerkin FEM, one of the biggest advantages of the least squares FEM is the resulting system of equations, which is always symmetric positive definite due to its origin from a minimisation problem (see also section 3.8). The ability to cope with first order terms without stabilisation is another advantage of the least squares FEM (cf. [83]). These advantages render the least squares FEM an interesting alternative for the solution of the Navier-Stokes equations. But some disadvantages of the least squares FEM should also be noted. From the practitioners point of view the most severe drawback is probably the lack of guaranteed conservation. While the FVM guarantees local conservation, and the FEM global conservation, the LSFEM can only guarantee conservation in the limit case as $h \to 0$. Especially the issue of mass conservation is subject of controversial discussions (cf.[49], [58], [22]). Another disadvantage is the increased number of unknowns, which also has a negative impact onto the matrix bandwidth through the larger element stencils. Depending on the used functional, the LSFEM often also has stronger regularity demands on the solution. This might not seem very important, but for some examples like the driven cavity, it could have a dramatic effect on the accuracy as shown in [12].

4.2 Current State of Research

In this section a sketch of the available literature about the least squares FEM for the Stokes and Navier-Stokes equations will be given. A more detailed study of the available literature can be found in [86]. The focus is put onto first order least squares formulations for the Stokes equations, which could also be used for the Navier-Stokes

equations. Formulations, which do not include the velocity as unknown (like the first order formulation proposed in [44], [88] and [42]) cannot be used for the Navier-Stokes equations due to the nonlinear convective term.

First the selected first order formulations for the Stokes equations will be presented with their specific properties. After that some results for the Navier-Stokes equations will be presented.

4.2.1 Incompressible Stokes Equations

The incompressible Stokes equations are:

$$-\nu_f \Delta \boldsymbol{v} + \nabla \tilde{p} = f \quad \text{in} \quad \Omega \tag{4.1}$$

$$\nabla \cdot \boldsymbol{v} = 0 \quad \text{in} \quad \Omega. \tag{4.2}$$

These equations must be accompanied by some boundary conditions, which are sometimes hard to find, if they are to correspond to a physical setting (one example is the outflow boundary condition, which is still subject of discussion, cf. [72]). In the mathematical analysis the following condition is often added:

$$\int_\Omega \tilde{p}\, d\Omega = 0. \tag{4.3}$$

It ensures a zero mean pressure and circumvents problems with the pressure unknown, which is only determined up to a constant by Eq. (4.1).

Currently several first order approaches, which include the velocity vector as unknown, have been used and/or analysed in literature:

- Velocity-Vorticity-Pressure (abbreviated as $up\omega$-formulation)
- Velocity-Stress-Pressure (abbreviated as uTp-formulation)
- Velocity-Velocity Flux-Pressure (abbreviated as uUp-formulation)

The abbreviations have been selected in accordance with the names of the unknowns in the original publications. Although these first order systems are equivalent in the continuous case, their mathematical properties differ. This will be shown in the next sections, where the main results for these equations will be summarised.

$up\omega$-Formulation

Introducing the vorticity $\omega = \nabla \times \boldsymbol{v}$ as an additional variable leads to the velocity-vorticity-pressure formulation (abbreviated as $up\omega$-formulation):

$$\begin{aligned} \nu_f \nabla \times \omega + \nabla \tilde{p} &= \tilde{\boldsymbol{f}} \quad \text{in} \quad \Omega && (4.4)\\ \nabla \cdot \omega &= 0 \quad \text{in} \quad \Omega && (4.5)\\ \omega - \nabla \times \boldsymbol{v} &= 0 \quad \text{in} \quad \Omega && (4.6)\\ \nabla \cdot \boldsymbol{v} &= 0 \quad \text{in} \quad \Omega && (4.7) \end{aligned}$$

In 2D the vorticity vector reduces to a single scalar, while in 3D the vorticity introduces three new unknowns. Thus Eq. (4.5) will only be present in the 3D case.

It is probably the most popular least squares formulation for the Stokes and Navier-Stokes equations. A large number of articles examine the mathematical (cf. [84], [34], [13], [15], [16], [17], [21] and more) and practical (cf. [121], [120], [122], [83], [85], [134] and more) properties of this formulation.

For the 2D case the ADN theory (cf. section 3.5 and [1]) can be applied directly to Eqs. (4.4-4.7), as the number of equations and unknowns is even. The 3D case is slightly more complicated and can be found in [21] or [83]. A first analysis was presented in [84]. Later Bochev and Gunzburger found out that the analysis was not completely correct (cf. [21]).

Depending on the principal part which is chosen for the analysis, the following two collections of a priori estimates which hold for positive indices q were found with the way described above ([21]):

$$\begin{aligned} \|\omega\|_{q+1,\Omega} + \|\tilde{p}\|_{q+1,\Omega} + \|\boldsymbol{v}\|_{q+1,\Omega} \leq \\ c_1(\|\nu_f \nabla \times \omega + \nabla \tilde{p}\|_{q,\Omega} + \|\nabla \times \boldsymbol{v} - \omega\|_{q,\Omega} + \|\nabla \cdot \boldsymbol{v}\|_{q,\Omega}) \end{aligned} \tag{4.8}$$

$$\begin{aligned} \|\omega\|_{q+1} + \|\tilde{p}\|_{q+1} + \|\boldsymbol{v}\|_{q+2} \leq \\ c_2(\|\nu_f \nabla \times \omega + \nabla \tilde{p}\|_{q,\Omega} + \|\nabla \times \boldsymbol{v} - \omega\|_{q+1,\Omega} + \|\nabla \cdot \boldsymbol{v}\|_{q+1,\Omega}) \end{aligned} \tag{4.9}$$

But each of these estimates holds only in conjunction with appropriate boundary conditions, which have to satisfy the complementing condition. A simple counterexample (cf. [21]) reveals that Eq. (4.8) does not hold with a a pure velocity boundary condition. Prescribing the pressure and the normal velocity would satisfy the complementing condition and thus lead to an optimally accurate method which requires only L^2 minimisation.

An alternative is the use of a formulation based on estimate Eq. (4.9) where the pure velocity boundary conditions satisfy the complementing condition. Unfortunately, this estimate requires for $q = 0$ the minimisation of H^1 norms, which can be done only by using C^1 finite element spaces with the already mentioned disadvantages (cf. section 3.2).

In the discrete case, weighted norms allow the replacement of the H^1 norm by a L^2 norm (cf. subsection 3.5.1), which leads to the following LSFEM functional:

$$\mathcal{J}(\omega, \boldsymbol{v}, \tilde{p}) = \frac{1}{2}\left(\|\nu_f \nabla \times \omega + \nabla\tilde{p} - \tilde{\boldsymbol{f}}\|_{0,\Omega}^2 + h^{-2}\|\nabla \times \boldsymbol{v} - \omega\|_{0,\Omega}^2 + h^{-2}\|\nabla \cdot \boldsymbol{v}\|_{0,\Omega}^2\right). \tag{4.10}$$

Because this functional uses different spaces than the ones appearing in Eq. (4.9), the functional is not coercive in the usual sense. But it is possible to establish the following error estimate which shows the convergence of the discretisation based on Eq. (4.10) (cf. [21], [19]):

$$\begin{aligned}\|\omega - \omega_h\|_{0,\Omega} + \|\tilde{p} - \tilde{p}_h\|_{0,\Omega} + \|\boldsymbol{v} - \boldsymbol{v}_h\|_{1,\Omega} \leq \\ c_1 h^k(\|\omega\|_{k,\Omega} + \|\tilde{p}\|_{k,\Omega} + \|\boldsymbol{v}\|_{k+1,\Omega})\end{aligned} \tag{4.11}$$

Here $k \geq 2$. To achieve optimality the polynomial degree of the ansatz-functions for the discretisation of $\boldsymbol{v}$ must be taken one order higher than for $\tilde{p}$ and ω.

Alternatively a negative norm, which corresponds to setting $q = -1$ in Eq. (4.9) can be used. This approach was proposed in [45] and leads to the following functional:

$$\mathcal{J}(\omega, \boldsymbol{v}, \tilde{p}) = \frac{1}{2}\left(\|\nu_f \nabla \times \omega + \nabla\tilde{p} - \tilde{\boldsymbol{f}}\|_{-1,\Omega}^2 + \|\nabla \times \boldsymbol{v} - \omega\|_{0,\Omega}^2 + \|\nabla \cdot \boldsymbol{v}\|_{0,\Omega}^2\right). \tag{4.12}$$

Again spaces with different polynomial degree are necessary to achieve optimal convergence rates.

uTp-Formulation

Introducing the stresses as new unknowns gives another popular first order formulation:

$$\sqrt{2\nu_f}\ \nabla \cdot \mathbf{T} - \nabla\tilde{p} = \boldsymbol{f} \quad \text{in} \quad \Omega \tag{4.13}$$

$$\nabla \cdot \boldsymbol{v} = 0 \quad \text{in} \quad \Omega \tag{4.14}$$

$$\mathbf{T} - \sqrt{2\nu_f}\ \mathbf{D}(\boldsymbol{v}) = 0 \quad \text{in} \quad \Omega. \tag{4.15}$$

Here $\mathbf{T} = \sqrt{2\nu_f}\mathbf{D}(\boldsymbol{v})$ denotes the stress tensor scaled by $(\sqrt{2\nu_f})^{-1}$ with $\mathbf{D}(\boldsymbol{v}) = (1/2)(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^T)$ being the symmetric part of the velocity gradient. In 2D the system has 6 equations and 6 unknowns and in 3D 10 equations and the same number of unknowns. Jiang points out that these variables are not independent due to the incompressibility condition. His conclusion is that the stress formulation actually has only 9 unknowns and equations. Nevertheless, the mathematical analysis of the formulation was done using the ADN theory in [17], and the main results can also be found in [19].

While the right boundary conditions can lead to H^1-coercivity in the $up\omega$-formulation, Eqs. (4.13-4.15) cannot be made elliptic in the sense of Petrovski (a definition can be

found on p.76 in [129]), and hence fail to be H^1-coercive(cf.[17]). But it is possible to define a functional similar to the one used for the formulation using the vorticity:

$$\mathcal{J}(\mathbf{T}, \boldsymbol{v}, \tilde{p}) = \frac{1}{2}(h^{-2}\|\mathbf{T} - \sqrt{2\nu_f}\mathbf{D}(\boldsymbol{v})\|^2_{0,\Omega} + h^{-2}\|\nabla \cdot \boldsymbol{v}\|^2_{0,\Omega} + \|\sqrt{2\nu_f}\nabla \cdot \mathbf{T} - \nabla\tilde{p} - \boldsymbol{f}\|^2_{0,\Omega}) \tag{4.16}$$

This functional is still not H^1-coercive, but makes it possible to establish optimal convergence results using weighted least squares functionals. Also the use of negative norm methods seems to be possible but was not examined yet.

One disadvantage of the velocity-stress-pressure formulation is the high number of equations and unknowns (6 Unknowns in 2D and 10 Unknowns in 3D), which induces higher computational costs without bringing significant advantages, if the stresses are not of interest.

According to the available literature, this formulation is rarely used for the Stokes equations. Most articles examine the extension of this formulation to the Navier-Stokes equations (cf. section 4.2.2).

uUp-Formulation

Another first order formulation was proposed by Chang ([47]), who called the method acceleration-pressure formulation. In this work the method will be called *uUp*-formulation despite the fact that the velocity gradient is denoted by $\mathbf{L}$ in this work. This name is inspired by the names of the unknowns in the works [14] and [20], where $\mathbf{L}$ was denoted by $\mathbf{U}$. The full velocity gradient tensor is introduced as a new unknown:

$$-\nu_f(\nabla^T\mathbf{L})^T + \nabla\tilde{p} = \boldsymbol{f} \quad \text{in} \quad \Omega \tag{4.17}$$

$$\nabla \cdot \boldsymbol{v} = 0 \quad \text{in} \quad \Omega \tag{4.18}$$

$$\mathbf{L} - \nabla\boldsymbol{v} = 0 \quad \text{in} \quad \Omega. \tag{4.19}$$

Obviously this makes the resulting system of equations significantly larger than the previously shown first order systems. In the 2D case the resulting system has 7 unknowns and in the 3D case already 13 unknowns. Nevertheless this first order system has some properties which render it an alternative to the other first order systems.

If the system Eqs. (4.17-4.19) is accompanied by a velocity boundary condition it is not H^1-coercive and hence the method minimising only the following functional:

$$\mathcal{J}(\boldsymbol{v}, \mathbf{L}, \tilde{p}) = \frac{1}{2}(\| - \nu_f(\nabla^T\mathbf{L})^T + \nabla\tilde{p} - \boldsymbol{f}\|^2_{0,\Omega} + \|\nabla \cdot \boldsymbol{v}\|^2_{0,\Omega} + \|\mathbf{L} - \nabla\boldsymbol{v}\|^2_{0,\Omega}) \tag{4.20}$$

is suboptimal (cf. [12]).

Two ways exists to establish again the desired coercivity results. First one might choose again different norms for the different parts of the equation, which results then

in the following functional:

$$\mathcal{J}(\boldsymbol{v}, \mathbf{L}, \tilde{p}) = \frac{1}{2}(\| - \nu_f(\nabla^T \mathbf{L})^T + \nabla\tilde{p} - \boldsymbol{f}\|^2_{-1,\Omega} + \|\nabla \cdot \boldsymbol{v}\|^2_{0,\Omega} + \|\mathbf{L} - \nabla\boldsymbol{v}\|^2_{0,\Omega}) \tag{4.21}$$

Again either weighted norms or a discrete H^{-1} method can be applied to this functional. To achieve optimal accuracy the ansatz-functions for $\boldsymbol{v}$ must also be one order higher than those for $\mathbf{L}$ and $\tilde{p}$.

The other way is to add some seemingly redundant equations, to obtain a system which is fully H^1-coercive. For the velocity-velocity flux-pressure formulation these equations are:

$$\nabla(\operatorname{tr}\mathbf{L}) = 0 \quad \text{in} \quad \Omega \tag{4.22}$$

$$\nabla \times \mathbf{L} = 0 \quad \text{in} \quad \Omega \tag{4.23}$$

$$\boldsymbol{n} \times \mathbf{L} = 0 \quad \text{on} \quad \Gamma. \tag{4.24}$$

(The first one describes the divergence gradient, which is obviously zero due to the fact that the field should be divergence free. In the second equation the symmetry of second derivatives is expressed in a compact manner ($\partial^2 v/\partial x \partial y = \partial^2 v/\partial y \partial x$)). The resulting functional is then:

$$\begin{aligned}\mathcal{J}(\boldsymbol{v}, \mathbf{L}, \tilde{p}) = \frac{1}{2}(\| - (\nabla^T \mathbf{L})^T + \nabla\tilde{p} - \boldsymbol{f}\|^2_{0,\Omega} + \|\nabla \cdot \boldsymbol{v}\|^2_{0,\Omega} + \|\mathbf{L} - \nabla\boldsymbol{v}\|^2_{0,\Omega} + \\ \|\nabla(\operatorname{tr}\mathbf{L})\|^2_{0,\Omega} + \|\nabla \times \mathbf{L}\|^2_{0,\Omega})\end{aligned} \tag{4.25}$$

Unfortunately, these additional equations lead to more equations than unknowns and therefore the system is not of ADN type anymore. But it is possible to show full H^1 coercivity for this pure L^2 functional. Hence the method achieves optimal accuracy for equal order interpolation of all unknowns (cf. [13], [12], [35]).

One drawback of this method is that the additional equations restrict the function spaces which are used to find a solution (cf. subsection 3.5.3). Hence instead of the original problem related spaces, the space $H^1 \times H^1 \times H^1$ is used to find a solution, which then represents only a projection of the exact solution onto this space. Numerical experiments performed with the counterpart of this formulation for the Navier-Stokes equations revealed that it has problems to converge to the correct solution in case of discontinuous boundary conditions (cf. [12]).

4.2.2 Navier-Stokes Equations

The presence of the nonlinear convective term in the Navier-Stokes equations significantly complicates the analysis. Therefore results currently exist only for the $up\omega$-formulation and the two variants of the uUp-formulation . These were obtained by adaptations of the abstract framework developed in [70]. A detailed description of this

framework would be beyond the scope of this work. Hence the presentation is limited to some theoretical and practical results for the most popular Navier-Stokes least squares formulations.

$up\omega$-Formulation

This first order formulation of the Navier-Stokes equations is probably one of the oldest and most used. It is considered in [121], [120], [122], [83], [21], [16], and many other publications. One advantage is the low number of additional unknowns (one for 2D problems and three for 3D problems) and the physical meaning of the vorticity, which is introduced as a new unknown.

Two ways could be considered to include the nonlinear convective terms into this first order formulation. Obviously it can be expressed with the velocity derivatives, which leads to:

$$\nu_f \nabla \times \omega + \boldsymbol{v} \nabla \boldsymbol{v} + \nabla \tilde{p} = \boldsymbol{f} \quad \text{in} \quad \Omega. \tag{4.26}$$

for the first part of the Navier-Stokes equations. Alternatively the vorticity can be used:

$$\nu_f \nabla \times \omega + \omega \times \boldsymbol{v} + \nabla \tilde{p} = \boldsymbol{f} \quad . \tag{4.27}$$

Which of these methods is the better one is not completely clear. In [21] and [16], Eq. (4.27) is used, while in [83] the first approach is suggested.

Similarly to the $up\omega$-formulation for the Stokes equations, using the pure L^2 minimisation leads to suboptimal convergence rates, if pure velocity boundary conditions are applied. This was first pointed out by Bochev in his PhD-Thesis (cf. [21]) and related papers (cf. [16]). A prior analysis of the convergence, performed by Chang et.al. (cf. [48]), was shown to be wrong.

Tang et.al. used the LSFEM for the simulation of instationary flows, often introducing some additional "difficulties". In [121] they solved the Navier-Stokes equations coupled with the thermal equations to simulate thermally driven flows. But they also verified the results for other benchmark problems like the lid driven cavity and the flow over an obstacle. Their results seem promising, but according to Gresho (Chapter 3.16.9, [72]) Tang admitted in a personal communication with Gresho that the method failed to get the right Strouhal number for a flow around a cylinder. According to Gresho the dissipative Euler backward method used by Tang could be responsible for that behaviour. In later articles Tang used the LSFEM in conjunction with the Crank Nicholson scheme for time integration, which is also the way proposed by Jiang for time accurate solutions. With this method Tang simulated a 3D lid driven cavity ([120]) and again thermocapillary flows in 2D and 3D ([122]).

uUp-Formulation

In [14] and [20], Bochev et.al. analyse the augmented and the negative norm uUp-formulation for the Navier-Stokes equations. The main focus of that paper lies in the error analysis of the nonlinear least squares functional.

For both formulations theoretically optimal preconditioners are proposed. While multigrid algorithms (cf. [73]) should work for the augmented formulation, a combination of a Laplace operator for the velocity field and a identity operator for the other unknowns is proposed for the negative norm variant (cf. [14]).

Although the analysis yields optimal error estimates, a computational test in [12] revealed severe problems with the augmented formulation, when the solution is not smooth enough. This observation is related to the high regularity demands ($\boldsymbol{u} \in H^2(\Omega)$) on the solution, which might be too strong for the Navier-Stokes equations.

Despite its interesting properties, only a few articles have examined the use of this formulation for real life problems on more complex domains. One exception are the works [76], [75] and [54] which use the augmented formulation for the solution of FSI problems.

uTp-Formulation

This formulation was used to compute approximate solutions for the Navier-Stokes equations in [63], [130], [7] and probably some more publications. A space-time version, including high order ansatz-functions, of this formulation was examined in [8]. The use of this formulation in conjunction with the p-Method has been proposed and analysed by Winterscheidt and Surana in [130].

While this formulation has been analysed for the Stokes equations in [17], currently their is no analysis available for the Navier-Stokes equations (cf. section 5.1 in [19])

4.2.3 Stokes- and Navier-Stokes Equations for Transient Flows

In Chapter 3 the two approaches for solving initial boundary values problems with the LSFEM were introduced. Especially the use of finite difference schemes was used to solve several example applications in 2- and 3 dimensions (cf. [121], [120], [122], [83]). Although the results presented therein seem very promising, a numerical analysis of the used schemes was not performed. Hence a theoretical foundation of the LSFEM for the Stokes- and Navier-Stokes equations for transient flows is still missing.

4.3 Implementation

In order to clarify the exact numerical procedures which were used for the computations shown in the subsequent sections, the general procedure will be shortly introduced.

First the linearisation which is performed before applying the variational principle to the differential operator will be explained. After that the temporal discretisation, which is also applied on the level of the strong form is explained. The resulting functionals and the bilinear forms are summarised in appendix B.

4.3.1 Linearisation

As already mentioned in section 3.6, in this work the method proposed by Jiang in [83] is used to treat the nonlinearity. Hence before applying the variational principle to the partial differential equation, the partial differential equation is linearised by using the first element of a Taylor series of the nonlinear operator around the current solution.

For the Navier-Stokes equations the nonlinear operator is:

$$\mathcal{F}(\boldsymbol{w}) = \begin{pmatrix} (\boldsymbol{v}\cdot\nabla\boldsymbol{v}) - \nu_f\Delta\boldsymbol{v} + \nabla\tilde{p} - \tilde{\boldsymbol{f}} \\ \nabla\cdot\boldsymbol{v} \end{pmatrix} \tag{4.28}$$

with $\boldsymbol{w} = (\boldsymbol{v}, \tilde{p})^T$. The operator $\mathcal{D}$ can be found by Eq. (3.17):

$$\mathcal{DF}(\boldsymbol{w})[\delta\boldsymbol{w}] = \begin{pmatrix} \delta\boldsymbol{v}\cdot\nabla\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\delta\boldsymbol{v} - \nu_f\Delta\delta\boldsymbol{v} + \nabla\delta p \\ \nabla\cdot\delta\boldsymbol{v} \end{pmatrix} \tag{4.29}$$

Then setting $\delta\boldsymbol{w} = \boldsymbol{w}_{n+1} - \boldsymbol{w}_n$, Eq. (3.89) for the Navier-Stokes can be obtained:

$$\begin{aligned} \boldsymbol{v}_{n+1}\nabla\boldsymbol{v}_n + \boldsymbol{v}_n\nabla\boldsymbol{v}_{n+1} - \nu_f\Delta\boldsymbol{v}_{n+1} + \nabla\tilde{p}_{n+1} &= \tilde{\boldsymbol{f}} + \boldsymbol{v}_n\nabla\boldsymbol{v}_n & (4.30)\\ \nabla\cdot\boldsymbol{v}_{n+1} &= 0 & (4.31) \end{aligned}$$

Eq. (4.30) is then the basis for the least squares methods. Starting with an initial guess $\boldsymbol{v}_0$, a first-order system which is equivalent to Eq. (4.30) is solved by a least squares formulation until convergence is reached. Numerically this approach can be troublesome in some cases as the convergence of the Newton method requires that the solutions of Eq. (4.30) are approximated with sufficient accuracy (cf. section 3.6).

4.3.2 Time Discretisation

When transient problem are to be solved with the least squares FEM, the usual approach is to replace the time derivative by a finite difference (cf. section 3.7). The Navier-Stokes equations for transient flows consist of the transient part $\mathcal{E}$ and a constraint $\mathcal{C}$. Then the abstract problem is

$$\begin{aligned} \frac{\partial\boldsymbol{v}}{\partial t} + \mathcal{E}(\boldsymbol{v},\tilde{p}) &= 0 & (4.32)\\ \mathcal{C}(\boldsymbol{v}) &= 0 & (4.33) \end{aligned}$$

with $\mathcal{E}(\boldsymbol{v},\tilde{p}) = \boldsymbol{v}\cdot\nabla\boldsymbol{v} - \nu_f\Delta\boldsymbol{v} + \nabla\tilde{p} - \tilde{\boldsymbol{f}}$ and $\mathcal{C}(\boldsymbol{v}) = \nabla\cdot\boldsymbol{v}$.

Now the time discrete version of Eqs. (4.32–4.33) becomes

$$\frac{\boldsymbol{v}_{n+1}-\boldsymbol{v}_n}{\Delta t}+\theta\mathcal{E}(\boldsymbol{v}_{n+1},\tilde{p}_{n+1})+(1-\theta)\mathcal{E}(\boldsymbol{v}_n,\tilde{p}_n) = 0 \tag{4.34}$$

$$\mathcal{C}(\boldsymbol{v}_{n+1}) = 0 \tag{4.35}$$

by using the class of θ-methods, and by setting $\theta = 1.0$ for the constraint $\mathcal{C}$. Here n denotes the time step which is related to the time by $t = n \cdot \Delta t + t_0$ where t_0 stands for the start time (usually $t_0 = 0$).

Computational tests showed that this canonical approach leads to strong temporal pressure oscillations when used with $\theta = 0.5$. Interestingly these oscillations did not appear in the average $\bar{p}_n = 1/2(\tilde{p}_n + \tilde{p}_{n+1})$. This led to the hypothesis that this behaviour comes from the constraint equations (i.e. the incompressibility constraint and the constraint defining the additional unknowns), which are imposed at $n+1$. A velocity field cannot satisfy Eq. (2.40) and Eq. (2.41) simultaneously. Therefore the pressure serves as a penalty unknown, which is related to the violation of exact incompressibility. Thus the pressure should be evaluated also at $n+1$. If imposing the incompressibility constraint at $n+1/2$ could cure the pressure oscillations as well was not examined. Thus the preferred and implemented approach uses the following slight variation:

$$\frac{\boldsymbol{v}_{n+1}-\boldsymbol{v}_n}{\Delta t}+\theta\mathcal{E}(\boldsymbol{v}_{n+1},\tilde{p}_{n+1})+(1-\theta)\mathcal{E}(\boldsymbol{v}_n,\tilde{p}_{n+1}) = 0 \tag{4.36}$$

$$\mathcal{C}(\boldsymbol{v}_{n+1}) = 0 \tag{4.37}$$

The resulting first order formulations are shown in appendix B.

4.4 Lack of Mass Conservation of LSFEM Formulations

One disadvantage appearing in many standard finite element methods for the incompressible Stokes and Navier-Stokes equations is the lack of local mass conservation (cf. [72]). Using special element combinations (cf. table 3.13 in [72]) or discretely divergence free finite elements will restore local mass conservation (cf. [126]). In the LSFEM even global mass conservation is violated in the currently available formulations, because the equation which ensures that the velocity field is divergence free is just another part in the minimised functional. Hence if the other equations dominate the residuum, the mass conservation can become quite weak. This is shown in [49] where a Stokes flow around a cylinder is simulated with the LSFEM. Although the author uses the weighted LSFEM functional Eq. (4.10), the mass conservation is violated at the parts where the channel is narrowed by the cylinder. The effect can be easily seen from the average of the velocity fields at the inflow and above and below of the cylinder. It comes out that more mass flows into the domain than goes by the cylinder, which clearly indicates a loss of mass.

Chang shows a way how to circumvent this disadvantage. He adds an additional constraint to the system of equations which ensures mass conservation in every triangular

finite element. The incompressibility constraint is then implemented using Lagrange multipliers. This allows the interpretation of his method as a kind of penalty formulation which unfortunately also leads to a saddle point problem. Thus the disadvantage of that method is the loss of some properties which originally motivated the use of the LSFEM.

Furthermore Deang and Gunzburger repeated similar numerical tests in [58], and were not able to confirm the findings of Chang. In their article several different combinations of boundary conditions and functionals are examined. To emphasise the mass conservation an additional weight was introduced into the equation which ensures mass conservation. Already small weights improved the mass conservation according to the cited article.

In [22] the mass conservation of some least squares formulations was examined. The results therein confirmed the findings of Chang. As those varying results do not give a clear answer about the quality of mass conservation of LSFEM methods, in the following paragraphs this issue will be analysed again using numerical experiments.

4.4.1 Narrowing Channel

The cited articles considered flows on fine grids around relatively complicated domains. This seems not very advantageous, as it was shown that the LSFEM is conservative in the limit case. Hence the finer grids may camouflage the problem of mass conservation.

For the numerical results which will be presented, a pretty simple setup was used. A channel with an adjustable narrowing in the middle, which is shown in Fig. 4.1, will be used to analyse the behaviour of the LSFEM with respect to the mass conservation.

A pure Stokes flow, with a kinematic viscosity of $\nu_f = 0.01$ is considered, as the effects regarding mass conservation will be similar for the Navier-Stokes equations. Because the uUp-formulation and the $up\omega$-formulation , including the negative norm variants, minimise the same norm $\|\nabla \cdot \boldsymbol{v}\|_{0,\Omega}$ to ensure mass conservation, it will suffice to examine one formulation. Here the $up\omega$-formulation will be used, as it has less unknowns.

At the inlet and the outlet a parabolic flow profile with a maximum horizontal velocity of 1 in the centre is prescribed and at the top and bottom wall a no-slip boundary condition is imposed. The pressure is prescribed at one node at the outflow to guarantee a uniquely defined pressure field. With these boundary conditions difficulties at the outflow boundary can be circumvented and global mass conservation is ensured. Furthermore these boundary conditions are valid for the weighted LSFEM and the Galerkin formulation.

For the analysis of the mass conservation, the mass flux at the narrowest or widest part of the channel is compared to the mass flux at the inflow. This allows the evaluation of the quality of local mass conservation.

To compare the differences and to identify the possible reasons for the difficulties of the LSFEM regarding mass conservation, the results obtained with the LSFEM will be compared with results obtained using a standard Galerkin variational scheme. Fortu-

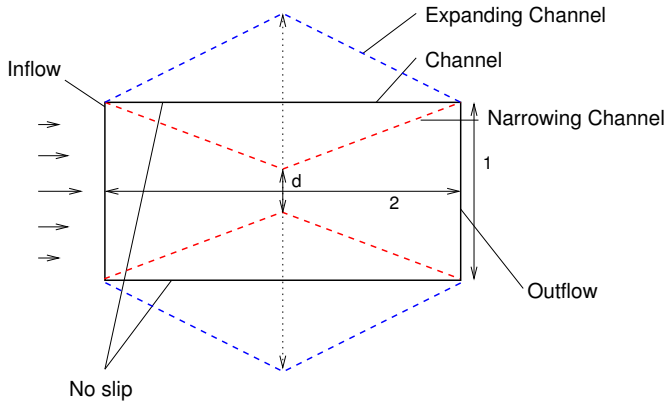

Figure 4.1. Setup of the channel flow with variable outlet diameter.

nately, the $Q_{k+1} - Q_k$ element pairs satisfy the inf-sup condition for $k \geq 1$ (for $k = 1$ this is the *Taylor-Hood* element cf.[72]. The proof of the inf-sup condition for $k \geq 2$ can be found in [28]), which allows a direct comparison with the LSFEM formulation.

In a first test, the geometry of the channel was modified over a range of parameters to get an impression of the influence of the geometry onto the mass conservation of the LSFEM and the Galerkin discretisation. Several grids, which were build by uniform refinement of the original grid consisting of two quadrilateral elements, were used.

A first observation, which can be seen from Fig. 4.2 is that the Galerkin as well as the least squares FEM fail to give sufficient mass conservation, when the channel becomes very narrow in the middle. The Galerkin FEM seems to perform slightly better with this geometry. A quite similar behaviour can be found on the other extreme. When the channel widens very strongly, the conservation properties of the Galerkin FEM seems to be slightly better. If the channel is straight, both methods find the exact solution, which can be exactly represented by the used $Q2 - Q1$ finite elements.

The weights for α_1^F and α_2^F were set to equal values, which were 1, 100 and 10000. Looking at the influence of the weight, which is shown in Fig. 4.3, it does not improve the mass conservation any more from a certain point on. It seems as if the found solution already is the minimiser of the L^2-norm of the divergence. But for finer meshes, the mass conservation can become worse, if the the weight is not chosen sufficiently high. This is in compliance with the predictions of the theory.

Comparing the L^2-norm of the divergence for the Galerkin and the least squares FEM indicates another problem (cf. Fig. 4.2). Although the Galerkin FEM achieves similar or better mass conservation, the L^2 norm of the divergence of the solutions found by it is always higher than that of the solution found by the LSFEM. Hence it seems as if the L^2-norm of the divergence is not a good measurement for the mass conservation. Although the conservative exact solution will obviously satisfy $\|\nabla \cdot \boldsymbol{v}\|_{0,\Omega} = 0$, in the

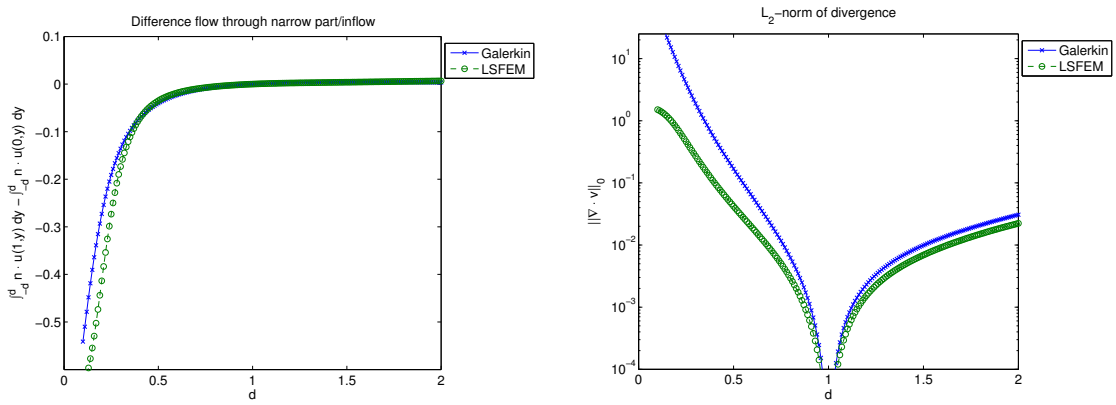

Figure 4.2. The difference between the mass going through the inflow and the narrowest part of the channel, the L^2 norm of the divergence over the whole computational domain. All results were computed on a $Q2 - Q1$ mesh with 32 elements. For the LSFEM the weights were $\alpha_1^F = \alpha_2^F = 100$ and $\alpha_0^F = 1$.

finite element spaces the function minimising this norm might not be the best with respect to the intended mass conservation.

Considering the geometry with a channel diameter of $D = 0.1$ at the narrowest part, a "reference" solution was computed using the 512 $Q5 - Q4$ elements with the Galerkin FEM. The highest velocity in horizontal direction was about 9. Using the LSFEM with a mesh consisting of 2048 $Q2 - Q1$ elements and setting $\alpha_1^F = \alpha_2^F = 10^6$, the solution found by the LSFEM shows a peak velocity of more than 17 and even parts with negative velocity (cf. Fig. 4.4). Thus arbitrarily increasing the weight of the term ensuring mass conservation can even lead to worse results!

Overall the observed effects could be attributed either to highly distorted elements or to an increasing strength of the corner singularity at the narrowing, which might have a negative effect on the accuracy of the LSFEM. Another benchmark problem will be used to obtain further information about this problem.

4.4.2 Porous Media Benchmark

A problem, where efficient solution strategies are not completely known yet, is the numerical simulation of flows through porous media. To test the quality of different codes, a benchmark problem was developed in cooperation with other institutes at the *Institute for Computer Applications in Civil Engineering* at the Technical University Braunschweig ([69]). It consists of several different geometries, which share a common property, namely a channel with several obstacles of square or cylindrical shape in it. The two geometries which were used for the following tests are shown in Figs. 4.5–4.6. Here this

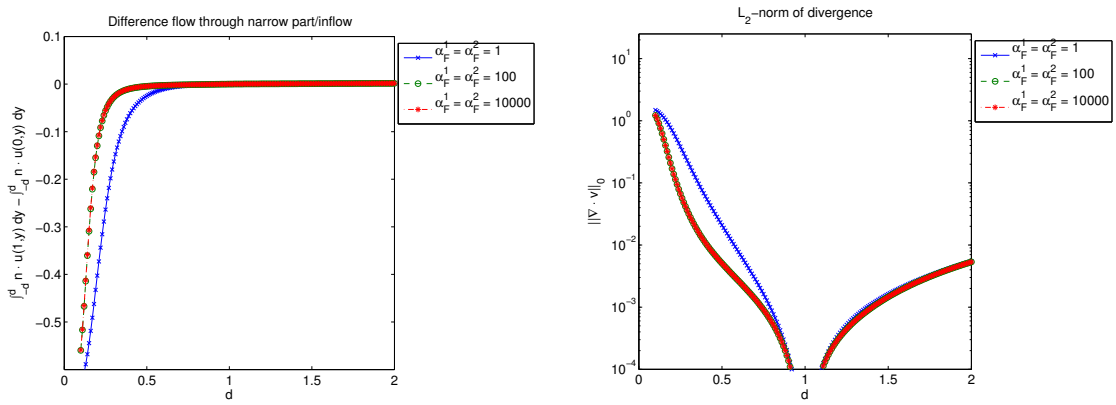

Figure 4.3. The difference between the mass going through the inflow and the narrowest part of the channel, the L^2 norm of the divergence over the whole computational domain. All results were computed with the LSFEM on a $Q2 - Q1$ mesh with 128 elements.

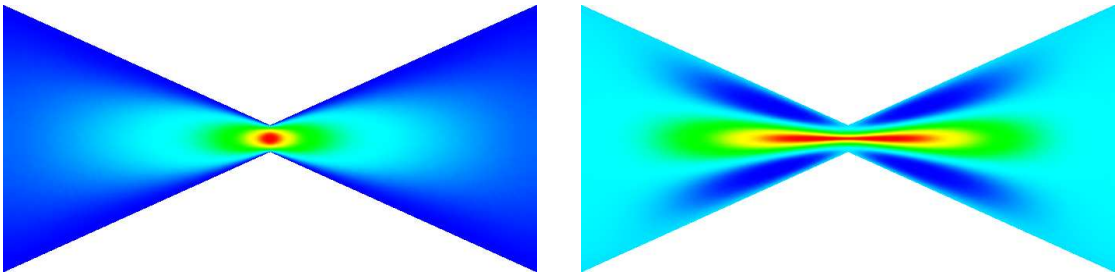

Figure 4.4. Horizontal velocity field (v_x). On the left a reasonable "reference" solution computed with the Galerkin method, using 512 $Q5 - Q4$ elements, and on the right the probably unphysical LSFEM solution, computed with 2048 $Q2 - Q1$ elements and $\alpha_1^F = \alpha_2^F = 10^6$, $\alpha_0^F = 1$.

benchmark will be used to analyse the mass conversation of the LSFEM in more complex domains.

Along the top and bottom wall and on the inflow, the velocity vector $\boldsymbol{v}$ is prescribed. On the boundaries of the obstacles, a no slip boundary condition is applied and at the outflow boundary the pressure is set to zero. The benchmark description only prescribes the Erguns Reynolds number which is defined by:

$$Re_E = \frac{|\boldsymbol{v}_0| D_p}{\nu_f \cdot (1 - \varepsilon_p)} \tag{4.38}$$

where D_p is the particle diameter, and the ε_p is a parameter, which is related to the porosity. In the benchmark problem, the particle diameter is $D_p = H/8$ with H being

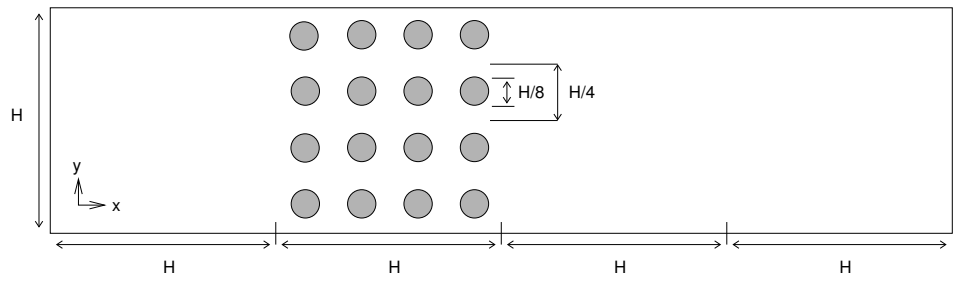

Figure 4.5. Porous media benchmark geometry with obstacles of cylindrical shape (Geometry I)

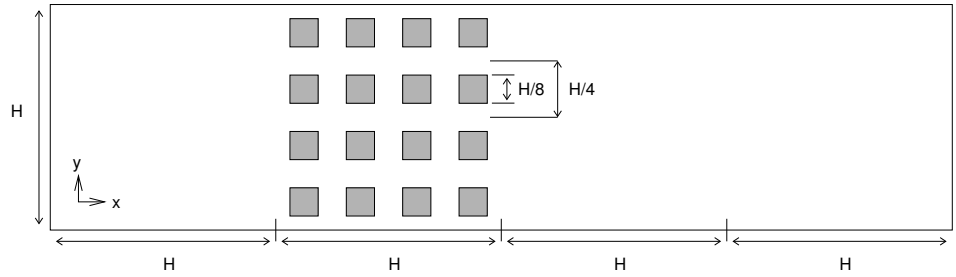

Figure 4.6. Porous media benchmark geometry with square shaped obstacles (Geometry II)

the channel height and the porosity is chosen to be $\varepsilon_p = 3/4$. Now to get the desired Reynolds number $Re_E = 1$, the relation between viscosity and inflow velocity is fixed. For an inflow velocity of $\boldsymbol{v}_0 = (1,0)^T$, the viscosity ν_f must be $1/2$.

Again the solution for this test problem was computed with several element pairs. In the geometry with the circular obstacles, the shape of the boundary was approximated using an sub/isoparametric mapping with the polynomial order of the shape functions for the pressure unknowns.

To visualise where the incompressibility constraint is violated, the divergence was computed pointwise on an overlay finite element grid with discontinuous elements. Furthermore the element L^2 norm of the residual of the incompressibility constraint was visualised element wise.

To get an idea of the general capability of the used $Q3-Q2$ Finite elements to represent a divergence free field in the given geometries, a pure Galerkin computation (stabilisation of the convective terms is not necessary due to the low Reynolds number) was performed on a mesh with 1280 elements for geometry I (cylinders) and with 1088 elements for geometry II (rectangles). The results for the two different geometries are shown in Figs. 4.7- 4.8. In both diagrams the same scale was used in Z-direction. Obviously there is a huge difference in the range of values for the two different geometries. This difference can be attributed to the edge singularities, which are much stronger in the domain with

the square shaped obstacles due to the shape of the obstacles. In the domain with the cylindrical obstacles, the singularities appear at the points where two element edges are connected. The numerical values of the divergence of the velocity field are between -28.5 and 28.2 for the geometry I and between -143.6 and 141.9 for geometry II.

Figure 4.7. $\nabla \cdot \boldsymbol{v}$ of the results obtained with a Galerkin computation with $Q3-Q2$ elements on the geometry I

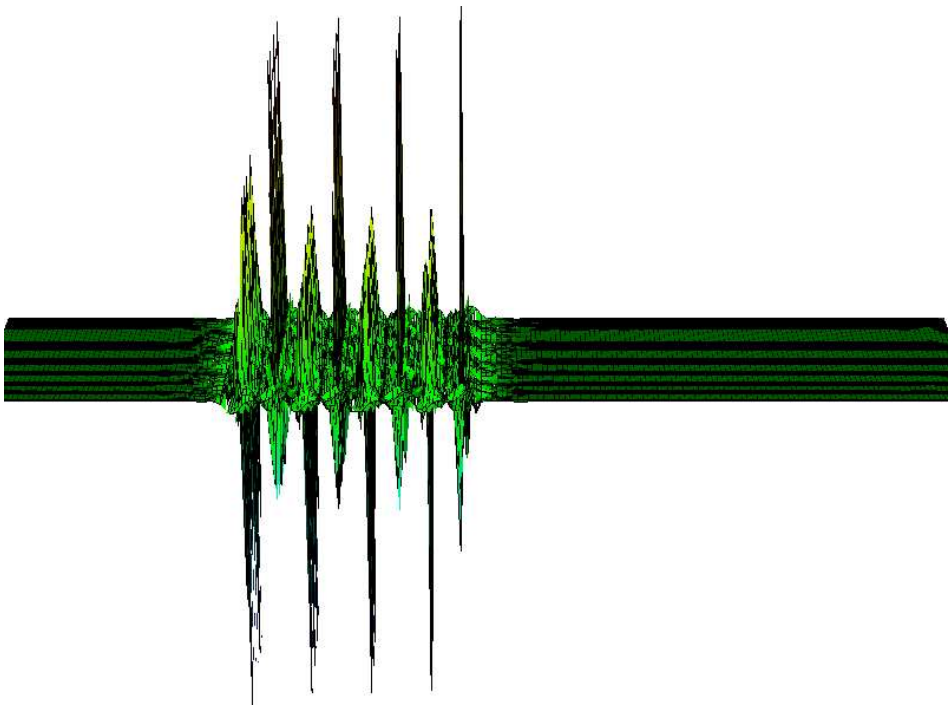

Figure 4.8. $\nabla \cdot \boldsymbol{v}$ of the results obtained with a Galerkin computation with $Q3-Q2$ elements on the geometry II

Hence a first finding is that the inevitable violation of the local mass conservation is strongest in the part of the domain with the obstacles, which is not really astonishing. Furthermore sharp corners lead to significantly higher local values of $\nabla \cdot \boldsymbol{v}$.

Now the results for the LSFEM calculations with weights $\alpha_1^F = \alpha_2^F = 40000$, using the same meshes and the same $Q3-Q2$ elements will be discussed. In Fig. 4.10 and Fig. 4.12 the velocity in horizontal direction is plotted. While the velocity field for geometry I shows approximately the correct behaviour, in geometry II, the LSFEM completely fails.

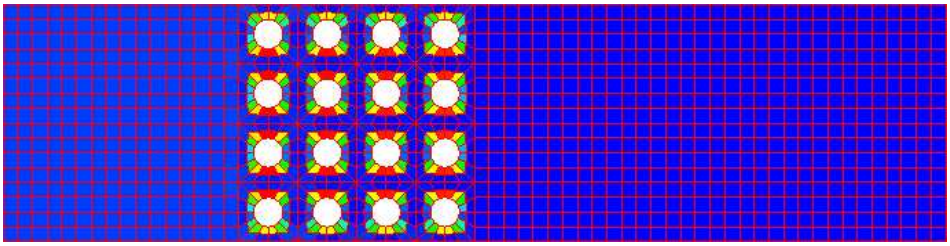

Figure 4.9. $\|\nabla \cdot \boldsymbol{v}\|_0$ on the elements, for a LSFEM computation with $Q3 - Q2$ elements on geometry I

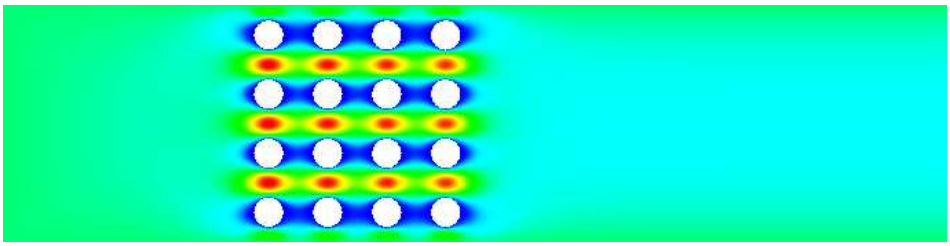

Figure 4.10. v_x for a LSFEM computation with $Q3 - Q2$ elements on geometry I

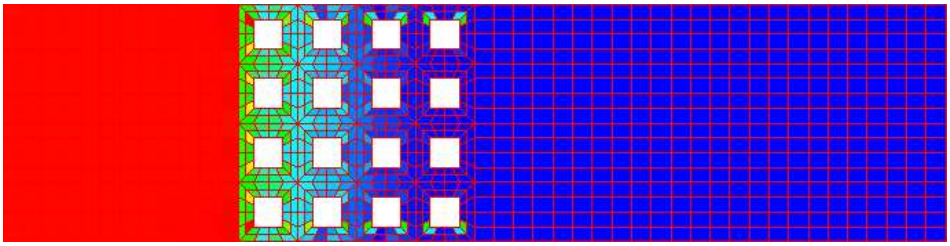

Figure 4.11. $\|\nabla \cdot \boldsymbol{v}\|_0$ on the elements, for a LSFEM computation with $Q3-Q2$ elements on geometry II

The distribution of the divergence term for geometry II, which is shown in Fig. 4.11, indicates that most of the mass is already lost in the inflow section. At a first glance, this behaviour seems very strange. Assuming a flow field where not most of the mass is not lost within the first section of the domain, much higher velocities would occur between the obstacles. These will also induce higher values of the divergence term especially in

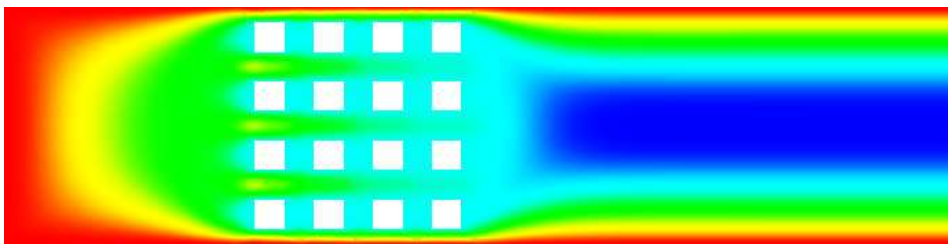

Figure 4.12. v_x for a LSFEM computation with $Q3 - Q2$ elements on geometry II

the corners. For geometry I, the L^2 norm of the divergence term is highest around the cylinders (cf. Fig. 4.9, which is similar to the Galerkin computation.

In the Galerkin method, the violation of the mass conservation is "weighted" linearly. By this we mean that having $\nabla \cdot \boldsymbol{v} = 10$ on an area of 1 unit, would count exactly as much as having $\nabla \cdot \boldsymbol{v} = 1$ on an area of 10 units. The LSFEM in contrast squares the absolute value of the divergence. Hence having $\nabla \cdot \boldsymbol{v} = 10$ in an area of 1 unit, would increase the minimised functional by exactly the same value as having $\nabla \cdot \boldsymbol{v} = 1$ on an area of 100 units!

Thus the solution minimising the functional lets the mass already disappear shortly after the inflow boundary conditions. Then in the rest of the domain, no significant mass loss appears and also violation of the conservation of momentum lead to lower values of the functional due to the lower absolute values.

4.4.3 Summary

Summarising the above findings, the LSFEM generally does not suffer from a lack of mass conservation. But the conducted numerical tests indicate that the results obtained with the LSFEM are very sensitive to sharp corners in the computational geometry. In these situations the minimisation of the L^2-norm of $\nabla \cdot \boldsymbol{v}$ is concentrated on the corner singularities and looses the overall picture. Furthermore increasing the weight of the L^2 term related to mass conservation can make the situation even worse. Here a sufficient weighting is necessary to obtain optimally accurate results, as already indicated by the theory. But using very high weights can be counterproductive. The Galerkin FEM seems to be more robust in these cases as it does not tend to overweight the corner singularities.

4.5 Analysis of Navier-Stokes Formulations for Transient Flows

As currently there are no theoretical results on the LSFEM solution of the Stokes and Navier-Stokes equations for transient flows, this section will be devoted to a basic computational analysis using the Taylor-Vortex, which is an exact solution of the Navier-Stokes equations for instationary flows on a square shaped 2D domain with periodic boundary conditions. On a domain $\Omega = [0,2] \times [0,2]$ the solution reads (cf. [72]):

$$v_x(x,y,t) = \pi \cos(\pi x) \sin(\pi y)\, e^{-2\pi^2 \nu_f t} \tag{4.39}$$

$$v_y(x,y,t) = -\pi \sin(\pi x) \cos(\pi y)\, e^{-2\pi^2 \nu_f t} \tag{4.40}$$

$$p(x,y,t) = \frac{1}{2}\pi^2 \left(\sin(\pi x)^2 + \sin(\pi y)^2\right) e^{-4\pi^2 \nu_f t} \tag{4.41}$$

with $\boldsymbol{v} = (v_x, v_y)^T$ and $\boldsymbol{x} = (x,y)^T$.

Although the problem of finding the numerical solution for this problem seems not complicated at a first glance, already small mistakes in the implementation lead to a loss of the full convergence order. Three setups will be used to test this formulation. All setups were tested with periodic boundary conditions, as severe problems were found in preliminary tests when setting the boundary conditions to the analytical solution on the boundary. Again the Galerkin formulation with $Q_{k+1} - Q_k$ elements will be analysed first. The details can be found on p. 438 of [109]. One of the least squares tests prescribed the periodic boundary conditions also on the vorticity field. Through this detail the bilinear form based on the non-weighted pure L^2 bilinear form (Eq. (B.32) with $\alpha_0^F \ldots \alpha_2^F = 1$) becomes fully H^1-coercive. This property should guarantee optimal convergence rates without weights and with equal order element pairs $Q_k - Q_k$. Finally the weighted formulation (Eq. (B.32) with $\alpha_0^F = 1, \alpha_1^F = \alpha_2^F = h^{-2}$) was tested without prescribing periodic boundary conditions for the vorticity.

Thus summarising this description, boundary conditions are:

$$\boldsymbol{v}((0,y)^T,t) = \boldsymbol{v}((2,y)^T,t) \quad \forall y \in (0,2) \quad \forall t \in (0,T] \tag{4.42}$$

$$\boldsymbol{v}((x,0)^T,t) = \boldsymbol{v}((x,2)^T,t) \quad \forall x \in (0,2) \quad \forall t \in (0,T] \tag{4.43}$$

for the Galerkin formulation and the weighted LSFEM formulation. The following additional boundary conditions were applied in the case of the unweighted LSFEM formulation:

$$\omega((0,y)^T,t) = \omega((2,y)^T,t) \quad \forall y \in [0,2] \quad \forall t \in (0,T] \tag{4.44}$$

$$\omega((x,0)^T,t) = \omega((x,2)^T,t) \quad \forall x \in [0,2] \quad \forall t \in (0,T] \tag{4.45}$$

On the corners of the domain the velocity was fixed by Dirichlet boundary conditions:

$$\boldsymbol{v}(\boldsymbol{x},t) = 0 \quad \text{at} \quad \boldsymbol{x} = \{(0,0)^T, (0,2)^T, (2,0)^T, (2,2)^T\} \quad \forall t \in (0,T] \tag{4.46}$$

Practically this means that the degrees of freedom belonging to the velocity in the corners are fixed to zero. While the Galerkin formulation did work properly without this additional constraint, the vortices started floating around in the domain without this additional constraint in the LSFEM.

The initial conditions $\boldsymbol{v}_h^0$ were defined by a L^2 projection. That is:

$$(\boldsymbol{v}_h - \boldsymbol{v}, \boldsymbol{w}_h)_{0,\Omega} = 0 \quad \forall \boldsymbol{w}_h \in \mathbf{V}_h(\Omega) \tag{4.47}$$

$$(\tilde{p}_h - \tilde{p}, q_h)_{0,\Omega} = 0 \quad \forall p_h \in \mathbf{Q}_h(\Omega) \tag{4.48}$$

$$(\omega_h - \omega, \kappa_h)_{0,\Omega} = 0 \quad \forall \kappa_h \in \mathbf{Q}_h(\Omega) \tag{4.49}$$

with $\boldsymbol{v}$ and p being the analytical solution defined by Eqs. (4.39–4.41) at $t = 0$. Eq. (4.49) is only required for the $up\omega$-formulation . The spaces $\mathbf{V}_h(\Omega)$ and $\mathbf{Q}_h(\Omega)$ are the used finite element spaces for the velocity and pressure or vorticity unknowns respectively.

To avoid the necessity for stabilisation in the Galerkin formulation the viscosity was set to $\nu_f = 0.1$ with a density of $\rho_f = 1.0$ which did not lead to observable stability problems due to oscillations.

4.5.1 Galerkin Formulation

For the computations the $Q5 - Q4$ element was selected, and a solution was computed on three different uniform refinements of the original discretisation consisting of one quadrilateral element. Fig. 4.13 shows the development of the L^2 error of the horizontal velocity v_x.

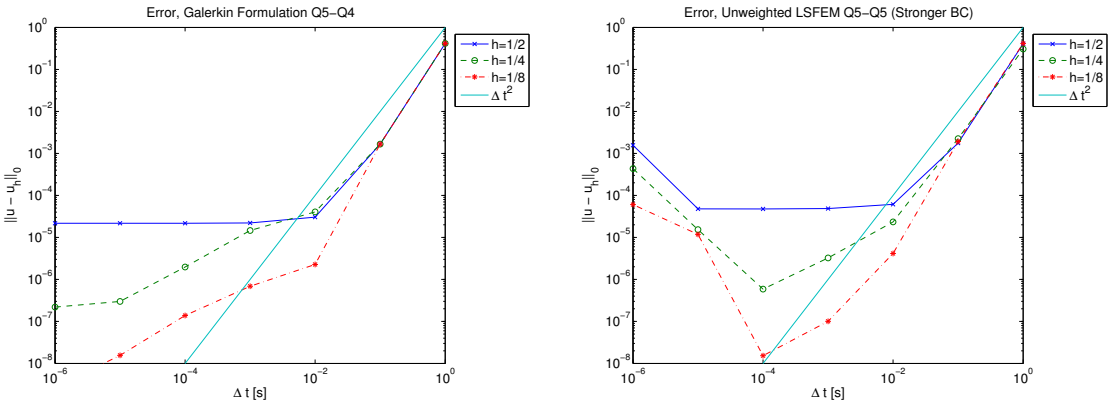

Figure 4.13. Results for the Galerkin formulation and the unweighted L^2 formulation with stronger boundary conditions (i.e. the periodicity of ω is prescribed)

The graph for $h = 1/2$ shows the expected behaviour. As long as the temporal error is dominant, the error decreases with second order. From a certain point on, the spatial error dominates the error and reducing the time step leads to no further error reduction.

Interestingly the curves for the finer meshes do not exhibit the same behaviour. When the time step size is reduced below some value, the error only decreases linearly with the time step size until again the spatial error becomes dominant. It is not clear what causes this behaviour. Possible reasons might be disturbances from the initial conditions or the convective terms. Nonetheless the Galerkin FEM leads to very accurate results with the $Q5 - Q4$ element and shows the expected second order convergence up to a certain Δt.

4.5.2 Unweighted $up\omega$-Formulation

The theory for the $up\omega$-formulation predicts that the pure L^2 formulation without mesh dependent weights (i.e. $\alpha_1^F = \alpha_2^F = 1$) becomes fully H^1-coercive if the right boundary conditions are selected (cf. [21]). Fig. 4.13 shows the results which were obtained with this formulation and the additional periodic boundary conditions Eqs. (4.44–4.45). The general convergence behaviour is comparable to that of the Galerkin formulation. For larger Δt the method achieves full second order accuracy, which is then limited by the spatial accuracy from a certain point on. Looking at the results for $\Delta t = 10^{-4}$ the full spatial accuracy of the used $Q5$ element is utilised. One uniform refinement leads to a reduction of the error to approximately one hundredth of the error on the coarser mesh. Theoretically it should be reduced by a factor of 64.

Making the timestep even smaller leads then to an increase in the error. This could be attributed to the used position of Δt within the formulation. As it is put in front of the finite difference term (cf. Eq. (B.31), Eq. (B.38)), the resulting system approaches the mass matrix as $\Delta t \to 0$. This leads to a reduced influence of the incompressibility constraint. Hence for small timesteps the distribution of Δt onto the different parts of the least squares functional has to be changed. Currently there are no articles available which consider this problem.

4.5.3 Weighted $up\omega$-Formulation

If the $up\omega$-formulation should be used with boundary conditions which do not suffice to guarantee the coercivity of the resulting bilinear form, the weighted approach has to be used (cf. [21]). This approach will be tested in this section. Two choices for the weights α_1^F and α_2^F will be examined. In the first test the weight remains unchanged and equal to one during mesh refinement while it is set to $\alpha_2^F = \alpha_1^F = h^{-2}$ in the second test. All tests were again performed with the $Q5 - Q4$ element.

Fig. 4.14 shows the results obtained with two different weightings. Although the results are influenced by the already mentioned effect of the somehow unbalanced weighting between the constraint and the time dependent part of the system of partial differential equations, the accuracy of the formulation with properly selected weights is clearly better. The general effect, that the error is first dominated by the temporal discretisation error and after passing a certain timestep Δt by the spatial discretisation error can again

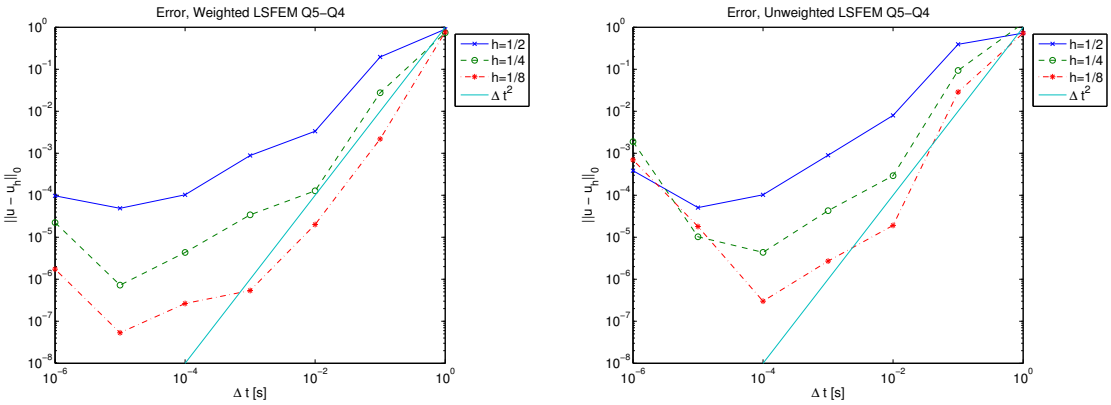

Figure 4.14. Comparison between $up\omega$-formulation without mesh dependent weights (right) and with mesh dependent weights α_1^F and α_2^F (left)

be seen in the diagram. Furthermore it confirms that the least squares FEM is second order accurate in time together with the θ-method and $\theta = 0.5$.

In a last computation the effect of the p-refinement was analysed on the finest spatial discretisation. The results are shown in Fig. 4.15 and demonstrate that the p-refinement works very well with the LSFEM. Here it should again be emphasised that the p-refinement for the Galerkin formulation was only possible through the comparatively high viscosity which stabilises the convective terms, while the LSFEM would work also if the viscosity were selected much smaller.

4.5.4 Summary

Three different formulation were computationally evaluated to test the accuracy in approximating solutions of the Navier-Stokes equations for transient incompressible flows. It was possible to confirm the second order accuracy in time which is predicted by the theory. For the $up\omega$-formulation the question of an optimal distribution of the different weights especially with respect to the influence of the time step size on the size of the residual is still open. Further research and a deeper mathematical analysis are clearly necessary.

4.6 Physical Weighting in the LSFEM

In the previous sections it was shown, that some formulations require a mesh dependent weighting, which can be justified by mathematical arguments. This section should examine the weighting with respect to the physical units used in the different parts of the equations.

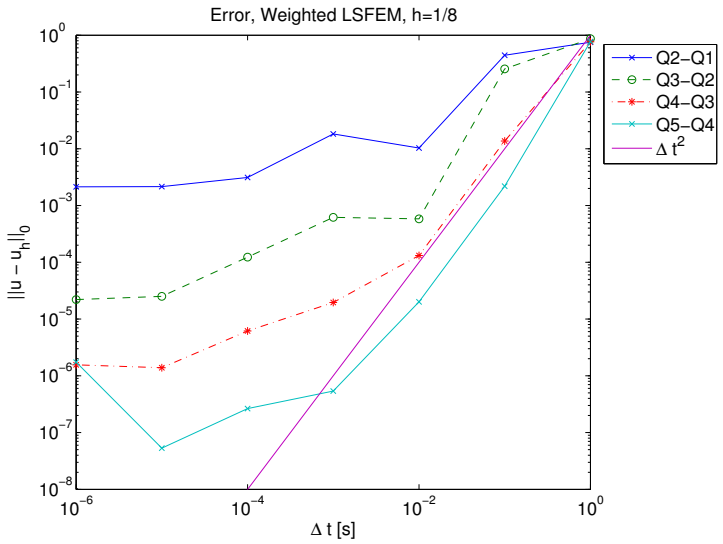

Figure 4.15. Effect of p-refinement onto the accuracy for the weighted $up\omega$-formulation

The LSFEM adds the squared norms of the equation residual and tries to minimise this value. Normally this would lead to the addition of different physical properties, which is wrong from a physical point of view, although mathematically the convergence is guaranteed. Another issue related to this topic is the fact, that different units could produce a different solution. Therefore it is necessary to have some weights, which would lead to the summation of the same physical quantities. This would then also imply independence from the used physical units.

All units will be in the SI-System. For the L^2-LSFEM it will be sufficient to consider the strong first order forms of the Stokes and Navier-Stokes equations. If the unit of the residual in the strong form is equivalent, the L^2 norm of the residual will obviously also have the same unit. First two characteristic parameters will be introduced:

$$\bar{\boldsymbol{v}} \quad \text{in} \quad \left[\frac{m}{s}\right], \qquad \bar{L} \quad \text{in} \quad [m]\,. \tag{4.50}$$

$\bar{\boldsymbol{v}}$ is the characteristic velocity, which could be chosen to be the average inflow velocity, and $\bar{L}$ is the characteristic length, which could be the channel diameter or the the size of an obstacle.

The two material parameters in the Navier-Stokes equations are the dynamic viscosity η_f and the density ρ_f which can be combined into one material parameter, the kinematic viscosity ν_f. These parameters have the following units:

$$\eta_f \left[\frac{kg}{ms}\right] \qquad \rho_f \left[\frac{kg}{m^3}\right] \qquad \nu_f = \frac{\eta_f}{\rho_f}\left[\frac{m^2}{s}\right] \tag{4.51}$$

The unit change caused by a one-time derivative corresponds to a division by the unit of the derivative direction. Thus for the spatial derivatives $\partial/\partial x$ and $\partial/\partial y$ a unit change of $[1/m]$ can be assumed.

Inserting the units into Eq. (2.40) shows that the residual of this equation has a unit of $[m/s^2]$. For Eq. (2.41), Eq. (4.6) and Eq. (4.19) the unit of the residual is $[1/s]$.

The weighted functionals Eq. (B.4) and Eq. (B.11) introduce the square of the characteristic element size as a weight for the L^2 norm. This squared length would correspond to weight of $[1/m]$ in front of the strong forms Eq. (2.41), Eq. (4.6) and Eq. (4.19). Hence introducing the following weights:

$$\alpha_0^F = 1, \quad \sqrt{\alpha_1^F} = \sqrt{\alpha_2^F} = \frac{\bar{L}\bar{\boldsymbol{v}}}{h} \left[\frac{m}{s}\right] \tag{4.52}$$

would lead to a functional with equivalent units satisfying the requirements of the stability estimate Eq. (4.11).

For the negative norm methods, the determination of physical weights is a bit more complicated. Here we will focus on the numerical procedure, which is used to evaluate the negative norm. As already explained, this process involves the computation of the inverse Laplace operator. The Laplace operator consists of the second spatial derivatives and hence maps a unit from [1] to $[1/m^2]$, following the same arguments as before. Thus the inverse Laplace operator will be a mapping from [1] to $[m^2]$. As

$$\|\cdot\|_{-1,\Omega}^2 = (\mathcal{S}\cdot, \cdot) \tag{4.53}$$

the unit of the squared residual will be

$$(\mathcal{S}[m/s^2], [m/s^2]) = ([m^3/s^2], [m/s^2]) = ([m^2/s^2], [m^2/s^2]). \tag{4.54}$$

Now using the following weights

$$\alpha_0^F = 1, \quad \sqrt{\alpha_1^F} = \sqrt{\alpha_2^F} = \bar{L}\bar{\boldsymbol{v}} \left[\frac{m^2}{s}\right] \tag{4.55}$$

would lead to an equivalent physical unit of the residuals in the functionals Eq. (B.6) and Eq. (B.13).

If real computations are intended, it is recommended to use these weights. Otherwise a change of the scales would have an influence onto the solution. This would clearly render the computed solutions useless for real life problems. Furthermore the inclusion of the mathematically required weights keeps the convergence rates optimal.

4.7 Evaluation of the Different LSFEM Formulations, Stationary Flow

In the following section, the different formulations for the Stokes and Navier-Stokes equations will be evaluated using a benchmark problem, which consists of a channel

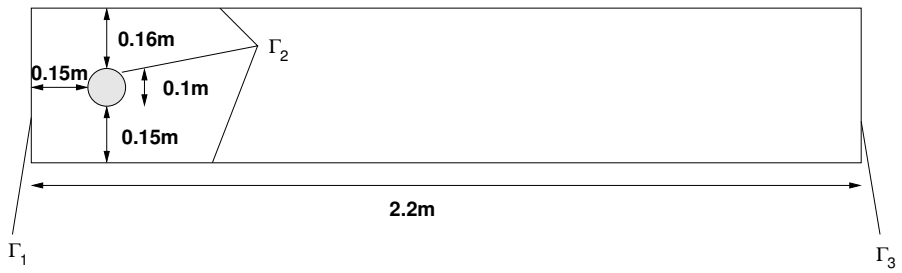

Figure 4.16. Scheme of the setup for the benchmark problem

flow around a cylinder, which is positioned slightly off the vertical centre. The setup is shown in Fig. 4.16. It was proposed by Schaefer and Turek in [114] to test different numerical schemes and computer codes for the solution of the incompressible Navier-Stokes equations. Hence a quite accurate reference solution is available.

The inflow velocity is described by a parabolic function, which leads to an average velocity of 2/3 the velocity in the centre of the channel. For the computation of the Reynolds number, the diameter of the cylinder $D = 0.1m$ is taken as characteristic length. With a kinematic viscosity ν_f of 1/1000, a velocity of $0.3m/s$ in the centre will result in $Re = 20$ and an inflow velocity of $1.5m/s$ will give a Reynolds number of $Re = 100$ (cf. [114] or p.6 in [126]).

At Reynolds number $Re = 20$ the flow field around the cylinder should settle at a stationary solution. Clearly the fluid will induce some drag onto the cylinder. Due to the asymmetric setup also a small lift should be noticeable. The difficulty for numerical methods mainly lies in the large difference between these two values, which differ about a factor of roughly 500.

Four important fluid quantities are used to evaluate the accuracy of the numerical method. These values are the lift and drag coefficients c_L and c_D, the pressure difference between the front and back of the cylinder Δp, and in the instationary case, the Strouhal number St. Lift and drag are defined according to the following equations:

$$c_D = \frac{2F_D}{\rho_f \bar{v}_f^2 D}, \qquad c_L = \frac{2F_L}{\rho_f \bar{v}_f^2 D}, \tag{4.56}$$

with

$$F_D = \int_{\Gamma_{cyl}} (\rho_f \nu_f \frac{\partial v_t}{\partial n} n_y - p\, n_x)\, d\Gamma \tag{4.57}$$

$$F_L = \int_{\Gamma_{cyl}} (\rho_f \nu_f \frac{\partial v_t}{\partial n} n_x + p\, n_y)\, d\Gamma. \tag{4.58}$$

In the following sections, different methods will be tested with respect to their accuracy. Therefore different element pairs are employed to study the influence of the

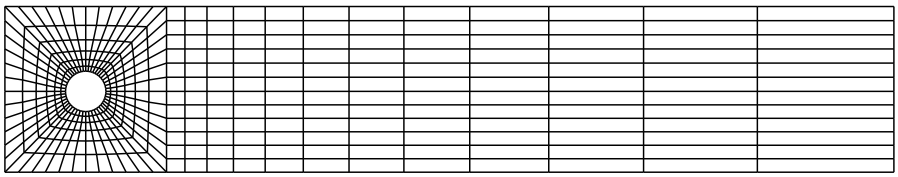

Figure 4.17. Spatial discretisation of the benchmark geometry

polynomial degree onto the accuracy. For each element pair, several computations are performed with different weights in the case of weighted formulations. For all setups, the lift- and drag coefficients are used as a measure for the accuracy.

4.7.1 Comparison of Galerkin FEM, L^2 LSFEM and H^{-1} LSFEM

As the convective terms in the Navier-Stokes equations require stabilisation in the Galerkin formulations and lead to degrading convergence rates for the negative norm LSFEM with the preconditioners proposed in [12] and [45], a first examination will use the Stokes equations to assess the quality of the different numerical methods.

For the computations the boundary was differentiated into three parts, the inflow Γ_1, the upper and lower channel wall and the cylinder boundary Γ_2 and the outflow Γ_3 (cf. Fig. 4.16). On these boundaries the following boundary conditions were prescribed:

$$\boldsymbol{v}(\boldsymbol{x},t) = (0.3(1-i(\boldsymbol{x},t)^2),0)^T \quad \text{on} \quad \Gamma_1 \tag{4.59}$$

$$\boldsymbol{v}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_2 \tag{4.60}$$

$$\tilde{p}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_3 \tag{4.61}$$

with $i(\boldsymbol{x},t) = (x_1 - 0.205)/0.205$.

For the kinematic viscosity a value of $\nu_f = 0.1$ was selected. With this setup the computations were done for several element pairs from $Q2-Q1$ to $Q5-Q4$ with the different methods on the spatial discretisation shown in Fig. 4.17. To utilise the full accuracy of the higher order elements, the edges along the cylinder boundary were approximated using a mapping with the lowest order appearing in the element pairs. Hence for the $Q3-Q2$ element the cylinder boundary was approximated with a $Q2$ mapping. A *reference solution* was computed on a triangulation which led to about 1 Mio. unknowns using the well verified code *Featflow* (cf. [126]).

Table 4.1 and Table 4.2 show the lift and drag coefficients for the different numerical methods. While the negative norm variants achieve an accuracy which is comparable to that of the Galerkin formulation for all element pairs, the L^2 methods do not work well for the lower order elements. With increasing polynomial degree the quality of the solutions found with the L^2 methods becomes much better with respect to c_L and c_D.

Whether the accuracy can be influenced by properly selected global weights will be analysed in the next sections for the $up\omega$-formulation and the uUp-formulation . But a

Element	Galerkin	$H^{-1} - up\omega$	$H^{-1} - uUp$	$L^2 - up\omega$	$L^2 - uUp$
Q2-Q1	2.9164	2.8340	2.7309	-0.5282	-0.2617
Q3-Q2	3.0133	2.9470	2.9744	2.8278	2.7933
Q4-Q3	3.0194	2.9729	3.0152	3.0166	3.0073
Q5-Q4	3.0196	-	-	3.0196	3.0190
Ref.	3.0273	3.0273	3.0273	3.0273	3.0273

Table 4.1. c_L on unrefined Grid

Element	Galerkin	$H^{-1} - up\omega$	$H^{-1} - uUp$	$L^2 - up\omega$	$L^2 - uUp$
Q2-Q1	307.57	307.01	291.81	71.31	61.62
Q3-Q2	313.57	313.96	311.63	303.02	293.18
Q4-Q3	314.20	314.20	314.08	314.04	312.82
Q5-Q4	314.24	-	-	314.24	314.14
Reference	313.89	313.89	313.89	313.89	313.89

Table 4.2. c_D on unrefined Grid

Element	Galerkin	$up\omega$	uUp
Q2-Q1	4188	4680	6156
Q3-Q2	9984	11832	17376
Q4-Q3	18372	22440	34644
Q5-Q4	29352	36504	57960

Table 4.3. Degrees of freedom in resulting algebraic system of equations

first result is that the negative norm methods offer better accuracy because they circumvent the difficulties related with the replacement of a stronger norm by a weighted L^2 norm. Nonetheless the Galerkin approach seems be the least problematic for the Stokes equations if the problems related to the solution of the resulting system of equations are neglected.

Looking at degrees of freedom, which are required by the discretisation of the finite element mesh with the used element pairs (cf. Table 4.3), the Galerkin variational scheme achieves clearly the most accuracy with respect to the number of degrees of freedom for a specific element pair. While the $up\omega$-formulation utilises only slightly more unknowns, the uUp-formulation increases the number of unknowns significantly.

4.7.2 $up\omega$-Formulation

As explained previously, the optimality of the $up\omega$-formulation depends on the right choice of mesh dependent weight which guarantees equivalence between the discrete H^1 norm and the discrete L^2 norm. The influence of the weighting parameter for this formu-

lation was examined in a very similar setting in [58]. In that article mainly the influence of the weighting onto the mass conservation was examined. Several procedures for the weighting were analysed. A global weight, which was determined by the average mesh size, an element by element weighting and a constant factor were used. Already global small positive weights (i.e. $\alpha_1^F > 5$) for the $\nabla \cdot \boldsymbol{v}$ term improved the mass conservation. Higher values for the weight did not lead to further improvements. According to the results presented in [58] also using local weights (i.e. choosing a different weight for each element in compliance with the size of that element) did not lead to different results.

The same boundary conditions as for the previous test which compared the least squares formulations for the Stokes flow, were use (i.e. Eqs. (4.59–4.61)). Again the element pairs $Q2-Q1$ – $Q4-Q3$ were tried out for the computations. For each element pair, the weights α_1^F and α_2^F were set to the same value which was varied over [1/8,16]. The third weight α_0^F is set to 1.

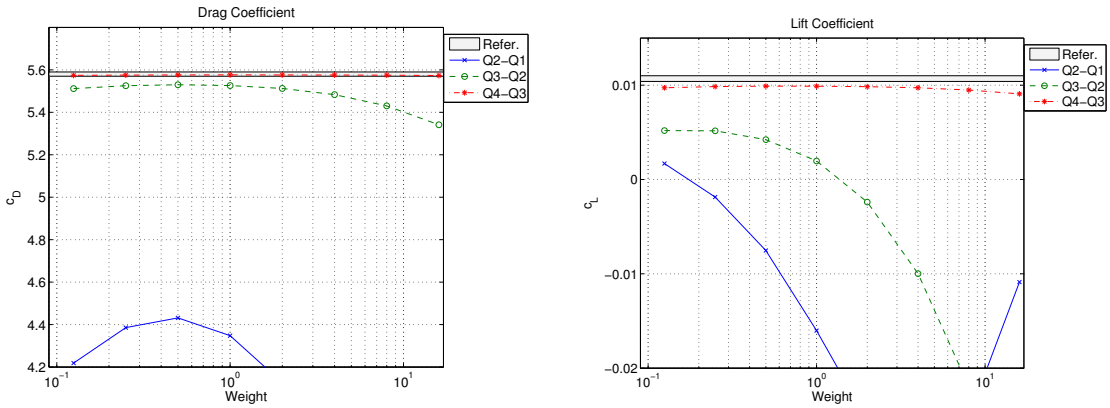

Figure 4.18. Drag and Lift coefficient for p-Refinement, $up\omega$-formulation

Fig. 4.18 shows the results for c_D and c_L, which clearly differ in dependence of the weight α_1^F and α_2^F. In that figure a grey bar marks the range of correct results. Mainly two results can be seen from Fig. 4.18. First the influence of the weight onto the accuracy decreases with increasing polynomial degree as the curves become flatter. The second interesting result is that the lower order element pair $Q2-Q1$ even with the seemingly best weight α_1^F and α_2^F is far from the correct result. Furthermore comparing the curves for c_D and c_L the best weights seems to differ depending on what parameter one is interested in. But the high order element $Q4-Q3$ seems to work properly and gives results which are in good agreement with the reference solutions while having a comparatively low dependence on the weight α_1^F and α_2^F.

The results for the h-refinement which are shown in Fig. 4.19 show a very similar behaviour. Increasing the approximation capabilities by refining the spatial discretisation widens the range of weights which give usable results. This should not be misunderstood

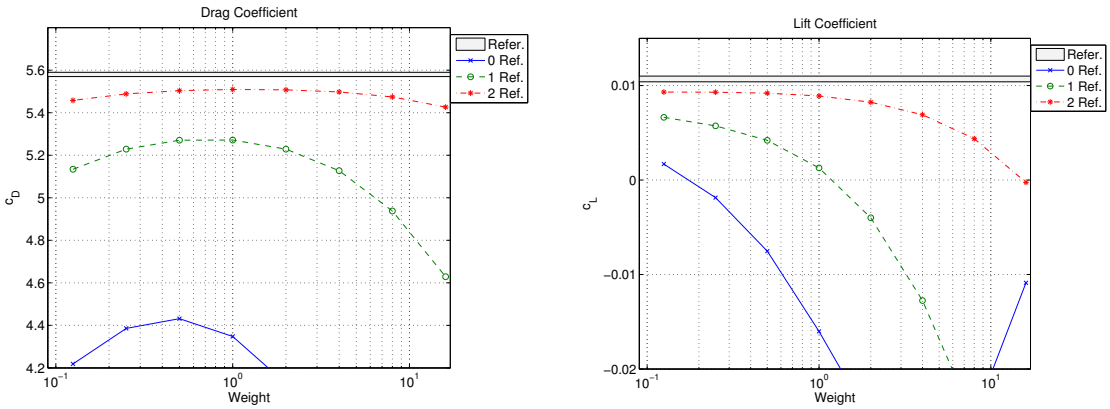

Figure 4.19. Drag and Lift coefficient for h-Refinement, $up\omega$-formulation

in the sense that the weight does not have to be increased during h-refinement. This can be seen from the curves for the drag coefficient in Fig. 4.19 where the maximum of the curves moves into the direction of higher weights for the finer discretisations.

These results show that the LSFEM works both as a h-method and as a p-method. From the practical point of view the p-method seems to be advantageous. Comparing the accuracy which can be achieved with the $Q2 - Q1$ element to the numerical costs measured in degrees of freedom, it is not very efficient compared to the same element used together with the Galerkin method. This can be seen in Table 4.1 and Table 4.2. Another issue is the required weighting which influences the condition number of the resulting systems of equations. In contrast the p-method works well for arbitrary p and does not lead to an increased condition number.

4.7.3 *uUp*-Formulation

The use of the full gradient tensor leads to the uUp-formulation , which could be employed either in the negative norm variant, the stabilised variant or with weighted norms, similar to the $up\omega$-formulation .

The boundary conditions are nearly the same as for the $up\omega$-formulation . These are mathematically admissible for the negative norm variant and hence should also be admissible for the weighted L^2 variant where the negative norm is replaced by a weighted L^2 norm. They are

$$\boldsymbol{v}(\boldsymbol{x},t) = (0.3(1-i(\boldsymbol{x},t)^2),0)^T \quad \text{on} \quad \Gamma_1 \tag{4.62}$$

$$\boldsymbol{v}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_2 \tag{4.63}$$

$$\tilde{p}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_3 \tag{4.64}$$

with $i(\boldsymbol{x}, t) = (x_1 - 0.205)/0.205$. No boundary conditions are imposed on the velocity gradient field U. The weights α_1^F and α_2^F are set to 1 while α_0^F is varied over [1/8,16].

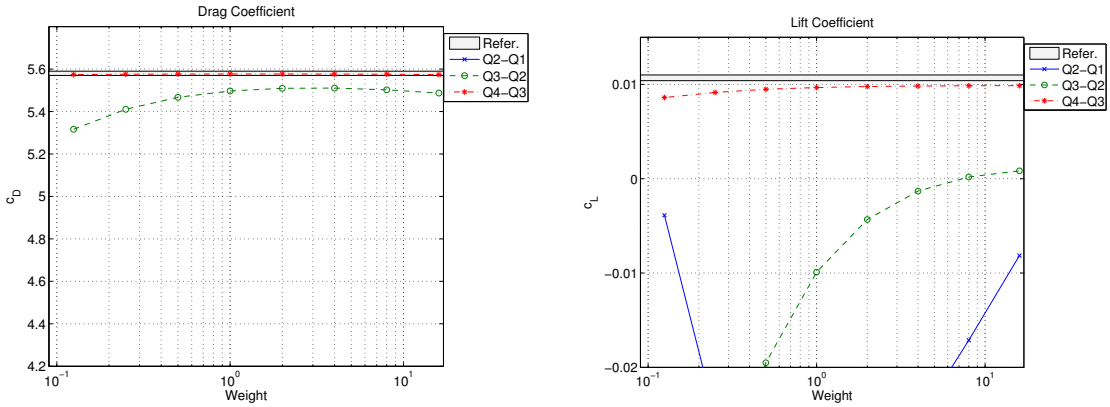

Figure 4.20. Drag and Lift coefficient for p-Refinement, uUp-formulation

Fig. 4.20 shows the results for c_D and c_L which were obtained by the uUp-formulation . Actually the qualitative behaviour is very similar to the one observed with the $up\omega$-formulation . The weight α_0^F has an influence on the accuracy but with the lower order element pairs the results always stay outside the admissible range on the coarse computational grid. Increasing the polynomial degree of the used shape functions does not only improve the accuracy but also reduces the dependence of c_D and c_L on the weight. This does not imply that the weight α_0^F can be ignored. A uniform refinement of the grid should be accompanied by an appropriate adaption of the weight.

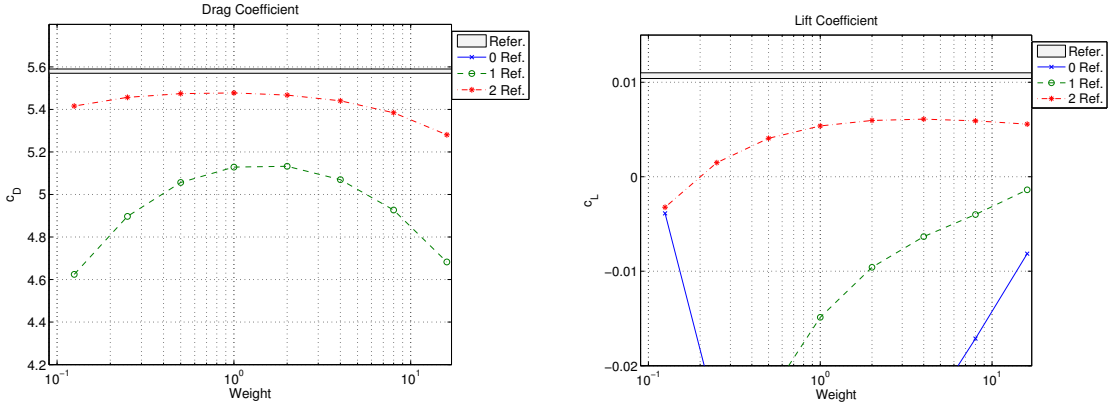

Figure 4.21. Drag and Lift coefficient for h-Refinement, uUp-formulation

This necessity can be seen in Fig. 4.21 which show the influence of uniform grid refinement on the lift and drag coefficient. As the weight α_0^F should compensate the differences

between the H^{-1} norm and the L^2 norm, the dependence between the characteristic element size h and the weight α_0^F is inversely to the relation for the $up\omega$-formulation where a stronger norm is replaced by the weighted L^2 norm. Thus the mesh refinement should be accompanied by reducing the weight α_0^F. In Fig. 4.21 it can be seen that the maximum of the curves for the drag moves into the direction of lower weights for reduced h.

These results show that also the weighed uUp-formulation is capable of finding the correct solutions to the benchmark problem. But the accuracy is comparable to that of the $up\omega$-formulation . Therefore the weighted L^2 variant of the uUp-formulation does not offer any significant advantage over the $up\omega$-formulation . Hence only the disadvantage of the higher number of unknowns remains.

4.8 Evaluation of the Different LSFEM Formulations, Transient Flow

This is again a test case from the benchmark problem, which was already used for the evaluation of the stationary Navier Stokes formulations (cf. [114]). An increased maximum inflow velocity of $1.5m/s$ gives a Reynolds number of $Re = 100$, which should lead to an instationary flow field. Actually the vortices behind the cylinder should separate from the cylinder and travel through the channel. In the temporal development of the lift and drag coefficients this behaviour should become visible through an oscillation. Again the difficulty is similar to that in the stationary test case, as the lift coefficient will undergo large oscillations, while the change in the drag coefficient is comparatively small and overlayed by a large permanent drag.

The instationary solution of the incompressible Navier-Stokes is again computed with different least squares schemes and different element pairs.

In addition to the temporal development of the lift and drag coefficients the Strouhal number $St = \frac{Df}{\bar{v}_f}$, which is a measure for the frequency of the vortex shedding, is determined. The shedding frequency f is determined from the temporal evolution of the lift coefficient by counting the timesteps between two zero passings. To increase accuracy several full oscillations in the quasistationary phase are used for this purpose.

4.8.1 $up\omega$-Formulation

Over a time interval of $5s$, the solution of the Navier-Stokes equations for transient flows was computed with a time step size of $\Delta t = 0.01$ and the finite difference approximation of the time derivative. The parameter θ in the time integration scheme was chosen to be 0.5. In each time step the nonlinear system of equations was solved with several Newton iterations. No simplifications or approximations, as proposed by Jiang in [83] were used.

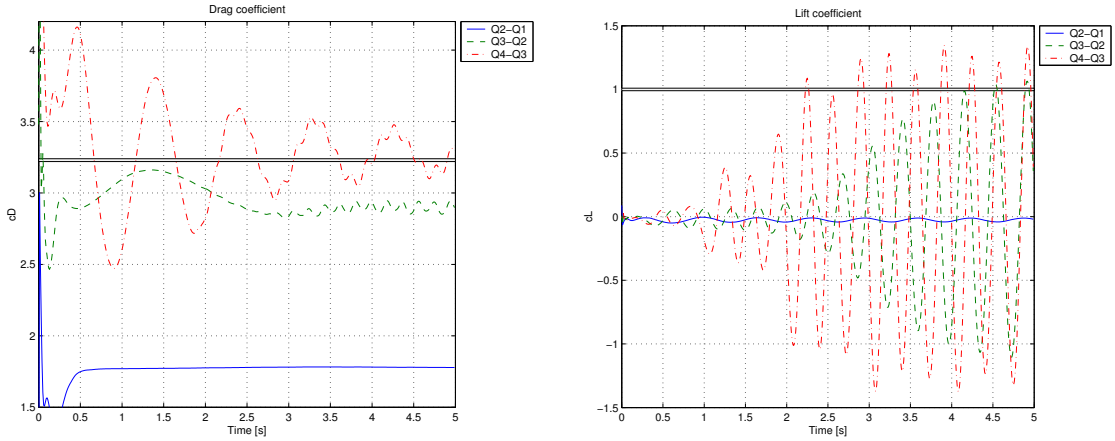

Figure 4.22. Drag and Lift coefficient, $up\omega$-formulation

Again the boundary conditions impose a parabolic inflow profile:

$$\boldsymbol{v}(\boldsymbol{x},t) = (1.5(1-i(\boldsymbol{x},t)^2),0)^T \quad \text{on} \quad \Gamma_1 \tag{4.65}$$

$$\boldsymbol{v}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_2 \tag{4.66}$$

$$\tilde{p}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_3 \tag{4.67}$$

with $i(\boldsymbol{x},t) = (x_1 - 0.205)/0.205$. At $t = 0$ the fluid was at rest ($\boldsymbol{v}(x) = 0 \quad \text{in} \quad \Omega$), and at $t > 0$ the full inflow velocity was prescribed. The weights were $\alpha_0^F = \alpha_1^F = \alpha_2^F = 1$.

The graphs in Fig. 4.22 show the development of the lift- and drag coefficient plotted over the time for the different element pairs. A grey bar denotes the interval, which was found to be acceptable in the benchmark. It can be clearly seen that the lower order element pairs $Q2 - Q1$ and $Q3 - Q2$ fail to find the correct values for the lift- and drag coefficients. The results for the stationary benchmark problem indicate that uniform h-refinement and adapted weighting would improve these results. Hence on a sufficiently refined spatial discretisation these element pairs would be usable as well.

But with the $Q4-Q3$ element pair, the formulation does produce very accurate results, and even the slight oscillation in the drag coefficient, can be seen. Currently the reason for the low frequency oscillation, which overlays the result, is not clear. It might be an artifact from the impulsive start and is clearly damped out, as time progresses.

For the lift coefficient, basically the same observations can be made, as for the drag coefficient. The lower order element pairs do not reach the right amplitude, while the $Q4-Q3$ pair achieves good results also for the lift coefficient. Beside the unphysical low frequency oscillation, the results are a bit too large.

Table 4.4 summarises the results for the Strouhal number. The results confirm the observations for the lift- and drag coefficients. Only the $Q4-Q3$ element lies within the interval from the benchmark.

Element	Strouhal nr.
Q2-Q1	0.1553
Q3-Q2	0.2703
Q4-Q3	0.2996
Low.lim.	0.295
Upp.lim.	0.305

Table 4.4. Strouhal number, $up\omega$-formulation

Element	Strouhal No.
Q2-Q1	0.1511
Q3-Q2	0.2756
Q4-Q3	0.3030
Low.lim.	0.295
Upp.lim.	0.305

Table 4.5. Strouhal number, uUp-formulation

For FSI problems, the Strouhal number is of great importance, because the instationary phenomena are often related to some resonance phenomena (cf. [127]). Thus the accurate prediction of shedding frequencies is an important demand for a fluid solver, which should be used for FSI problems.

4.8.2 *uUp*-Formulation

For the uUp-formulation the numerical tests as for the $up\omega$-formulation were performed. The time step size, the computational grid and the boundary conditions were the same as for the $up\omega$-formulation (Eqs. (4.65–4.67)). Thus the boundary conditions for $\mathbf{L}$ were not prescribed anywhere. Again different element pairs were used for the computations. The weights were $\alpha_0^F = \alpha_1^F = \alpha_2^F = 1$.

The results for the Strouhal number can be found in Table 4.5 and indicate a similar behaviour as found for the $up\omega$-formulation . With the high order elements, the method performs reasonably, while the lower order elements do not give satisfactory results, at least on the coarse grid.

The temporal evolution of the lift and drag coefficient is shown in Fig. 4.23. It is also quantitatively and qualitatively very close to the results obtained with the $up\omega$-formulation .

An coarse estimate of the required numerical effort can be obtained from Table 4.3 which shows the number of unknowns for the different formulations. As the results with the uUp-formulation are not significantly better than those which were computed with the $up\omega$-formulation , clearly the higher number of unknowns is a disadvantage of the

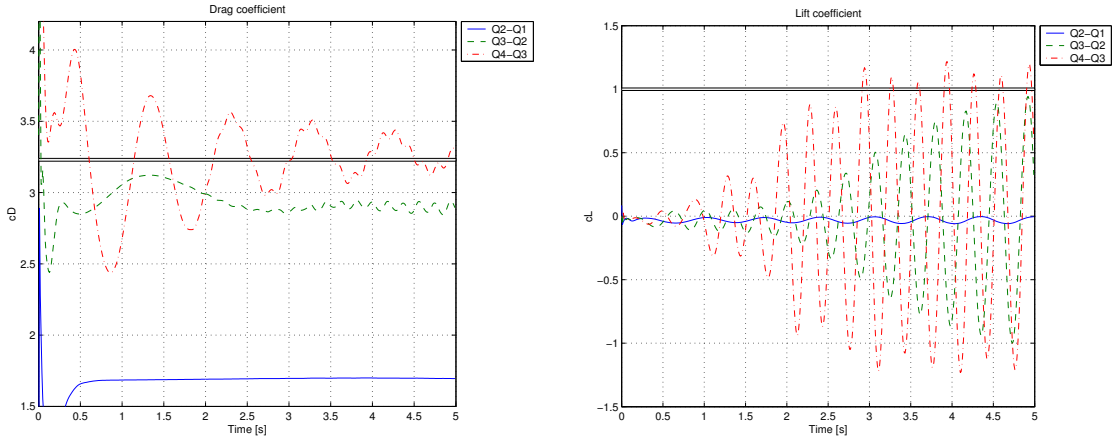

Figure 4.23. Drag and Lift coefficient, uUp-formulation

uUp-method and renders at least the weighted formulation more or less useless due to the missing benefits over the $up\omega$-formulation .

4.9 Geometric Conservation

Besides the physical conservation laws, the numerical procedures for the simulation of fluids in moving domains must satisfy another criterion, the geometrical conservation law (GCL). It simply demands that a uniform flow is invariant against grid motions. If the GCL is fulfilled by a numerical procedure, the stability properties of the non ALE procedure carry over to the ALE variant (cf. [64]). The GCL requires that the order of the time integration scheme is sufficient to exactly integrate the grid motion.

In [108] the GCL is derived for Galerkin methods with isoparametric finite elements. The article shows that the used time integration scheme must integrate polynomials of degree $kd-1$ exactly to satisfy the GCL. Here k is the polynomial degree of the mapping and d is the spatial dimension. It is not clear, if this criterion also holds for least squares methods.

As the used finite elements are independent of the used variational principle, that part of the theory needs no changes. But as it was already explained, the accuracy of the time integration is connected to the spatial discretisation. The literature does not provide any results on this topic. Thus the issue of geometric conservation will be examined experimentally by some computations.

For this purpose the temporal evolution of a uniform flow on a unit square will be examined:

$$\boldsymbol{v}(\boldsymbol{x},t) = (\sin(\varphi),\cos(\varphi))^T \quad \text{in} \quad \Omega \tag{4.68}$$

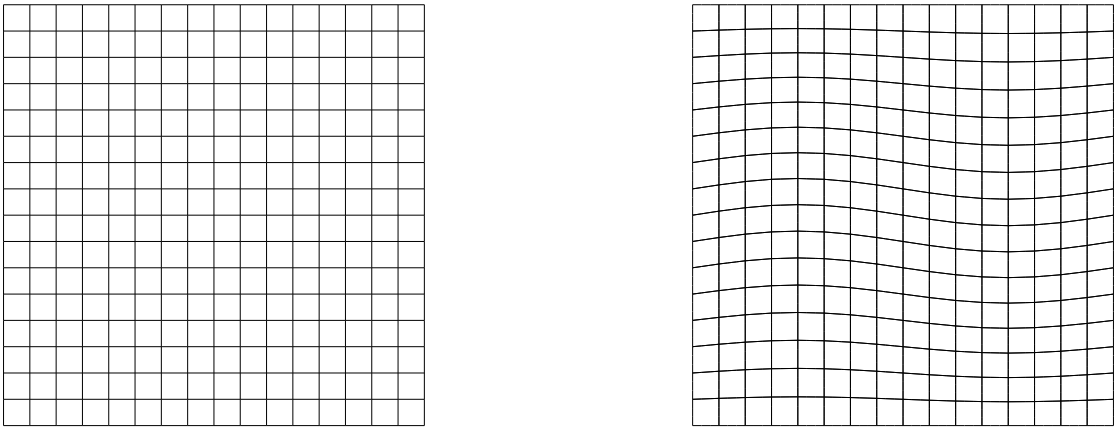

Figure 4.24. Initial and deformed fluid mesh at $t = 0.1$.

φ	$\|v_x - v_{h_x}\|_0$	$\|v_y - v_{h_y}\|_0$	$\|\tilde{p} - \tilde{p}_h\|_0$	$\|\omega - \omega_h\|_0$
0	$5.522922 \cdot 10^{-16}$	$2.959438 \cdot 10^{-15}$	$2.061750 \cdot 10^{-14}$	$1.203658 \cdot 10^{-14}$
0.25π	$6.944082 \cdot 10^{-15}$	$2.694506 \cdot 10^{-15}$	$9.403922 \cdot 10^{-14}$	$3.433702 \cdot 10^{-14}$
0.5π	$1.575347 \cdot 10^{-15}$	$2.141432 \cdot 10^{-15}$	$2.691865 \cdot 10^{-14}$	$1.506261 \cdot 10^{-14}$
0.75π	$2.244708 \cdot 10^{-15}$	$6.696721 \cdot 10^{-16}$	$4.378524 \cdot 10^{-14}$	$1.276208 \cdot 10^{-14}$
1π	$3.479516 \cdot 10^{-15}$	$2.454059 \cdot 10^{-15}$	$5.352162 \cdot 10^{-14}$	$3.025933 \cdot 10^{-14}$
1.25π	$2.139563 \cdot 10^{-15}$	$2.538200 \cdot 10^{-15}$	$4.768353 \cdot 10^{-14}$	$2.354982 \cdot 10^{-14}$
1.5π	$4.048065 \cdot 10^{-15}$	$1.311796 \cdot 10^{-15}$	$6.998472 \cdot 10^{-14}$	$2.277552 \cdot 10^{-14}$
1.75π	$4.719171 \cdot 10^{-15}$	$4.042928 \cdot 10^{-15}$	$8.390074 \cdot 10^{-14}$	$3.161298 \cdot 10^{-14}$

Table 4.6. Error in the problem variables for some flow directions φ.

The angle of the flow direction can be defined by φ. All weights $\alpha_0^F \dots \alpha_2^F$ were equal to one. On all boundaries the velocity is prescribed and at one node the pressure is fixed to get a uniquely defined pressure field. If the formulation satisfies geometric conservation, the flow field should not change when the mesh is moved. The mesh velocity was prescribed by:

$$\boldsymbol{v}_\Phi(\boldsymbol{x}, t) = (0, (y - y^2)\sin(2\pi x) \quad \text{in} \quad \Omega \tag{4.69}$$

After one timestep with the ALE version of the $up\omega$-formulation for the Navier-Stokes equations (i.e. the method based on the bilinear form Eq. (B.32)) the deviation from the exact solution in the velocity field was computed in the L^2 norm. Fig. 4.24 shows the undeformed mesh at $t = 0$ and the deformed mesh at $t = 0.1$ respectively.

The error after one timestep $\Delta t = 0.1$ is shown in Table 4.6. For different angles φ the size of the error is in the range of the floating point accuracy which leads to the conclusion that the least squares FEM satisfies the GCL for $\theta = 0.5$ and a Q1-mapping.

It should be noted that the fulfilment of the GCL does not imply that any flow is invariant against mesh motion. If a nonuniform flow field is given on a finite element mesh, moving the mesh through the field is equivalent to transporting the field through the mesh, which can be easily seen from the ALE equations. Thus in the Galerkin case the ALE convection has to be stabilised in a similar fashion as other convective terms. This is not necessary in the LSFEM due to its inherent stability regarding first order terms. But the transport of an arbitrary flow field nonetheless introduces numerical errors, which lead to unphysical disturbances. This is not specific to the LSFEM but generally related to the ALE formulation independently of the variational principle used.

4.10 Summary

After an overview about the existing literature for first order least squares finite element methods for the incompressible Navier-Stokes equations, the implementational details were discussed. Two of the many first order formulations which were proposed so far were selected for further investigation. The $up\omega$-formulation as the most popular formulation which also was properly analysed mathematically, and the uUp-formulation which is a relatively new development but looks very promising.

Following the idea presented in [83] the Navier-Stokes equations were already linearised in their strong form by using the first terms of a Taylor series around the current solution. This procedure leads to a system of linear partial differential equations which, starting with an initial guess, give the solution of the next Newton iteration.

A critical aspect of the least squares methods is their lack of global mass conservation which is a widely demanded minimal requirement for numerical procedures for the incompressible Stokes or Navier-Stokes equations. Hence this problem was numerically analysed with two new problem setups. Especially the porous media test revealed that the criticised lack of mass conservation is highly problem dependent. In domains with smooth boundaries the achieved mass conservation has led to reasonable results, while in a very similar setup with nonsmooth boundaries the mass conservation was so poor that the result was completely wrong. A comparison with results obtained using a Galerkin formulation for the Stokes equations has led to the hypothesis that this effect might be attributed to the different weighting of a violation of mass conservation, which makes the LSFEM more sensitive to singularities.

While there exists some literature which demonstrates the use of the least square FEM for the Navier-Stokes equations describing transient flows, no deeper mathematical analysis of these procedures has been performed yet. This motivated a computational analysis of the least squares methods for the Navier-Stokes equations describing transient flows. The results showed that the LSFEM obtains optimal spatial and temporal discretisation errors. But the results also indicate that the weights have to be adapted if the time step size is changed. Theoretical results for other transient problems (cf.

[118]) confirm this finding. How the weights have to be changed exactly remains an open question.

To evaluate the accuracy of the different first order formulations, another benchmark problem which has been employed by many authors was used. Here the focus was shifted away from the more fundamental characteristics of the least squares methods to the properties which are relevant for practical applications. For the Stokes equations the solutions found by the uUp-formulation and the $up\omega$-formulation in their negative norm variant and the weighted L^2 variant have been compared with a solution obtained by the well verified code Featflow and a Galerkin formulation with inf-sup stable elements. Due to the real H^1 coercivity, the results of the negative norm variants were better than those obtained with the L^2 variants. Increasing the polynomial degree of the finite element shape functions reduced this effect and led to accurate results even for the weighted L^2 formulations.

Due to difficulties with the convective terms the tests for the Navier-Stokes equations were only performed with the weighted L^2 variants. The general behaviour of the uUp-formulation and the $up\omega$-formulation was nearly equivalent. With higher order finite elements the influence of the weighting parameter decreased, while the lower order elements did not give accurate results on the finite element discretisation used.

The analysis of the least squares methods for transient flow problems showed very similar results. Only the $Q4 - Q3$ element gave satisfying results on the used computational grid independent of the first order formulation used. For this element pair also the found Strouhal number fitted well into the interval of admissible solutions defined by the benchmark description.

To conclude this summary the numerical tests showed that the least squares FEM is capable of finding accurate solutions for the Stokes and Navier-Stokes equations describing stationary and transient flows under some restrictions. Nonsmooth boundaries could lead to degenerated solutions and also false weighting could have a negative effect on the accuracy. Especially in the transient case the proper weights are not clear yet because a sound mathematical theory is missing. But for higher order elements the influence of the weight decreases. Therefore these high order elements will be used for the fluid structure computations as they currently seem to be the most reliable. The negative norm methods could be an option, but the question of efficient preconditioning in the case of significant convection has not yet been resolved. As the uUp-formulation does not offer any advantages in conjunction with the L^2 LSFEM, the $up\omega$-formulation is currently preferred due to the lower number of unknowns.

5 Numerical Methods for the Structural Domain

Similar to chapter 4 first a short overview about the general approaches for the equations of linear elasticity will be given. After that the current least squares approaches for these equations will be summarised. Then a short section about the implementation of a numerical method based on the Galerkin variational principle follows. This numerical method will be used for LSFEM-Galerkin coupled solution algorithms for FSI-problems and for the verification of the least squares formulations, which will be developed in the subsequent sections.

As literature which considers least squares methods for the solution of transient problems is rare, some fundamental results for the least squares method are presented and illustrated using the wave equation as an example. The general idea is then used to analyse three new least squares formulations for the solution of the equations describing transient linear elasticity problems. To evaluate the properties of these new formulations, numerical results are presented and discussed afterwards. This chapter is then closed with a short summary of the results found in the preceding sections.

5.1 Numerical Procedures for the Equations of Elasticity

The origins of the finite element method are in the area of structural problems, where they were first used. Static structural problems require the minimisation of the potential energy. Therefore the Galerkin variational principle, which works very well for this type of problem, leads to a symmetric positive definite system of equations and optimal convergence rates can be shown mathematically. Problems occur if the material becomes nearly incompressible. In this case the standard finite element basis shows *locking* (cf. [27], [137]), which leads to degrading solution accuracy. Another disadvantage of the unmodified standard finite element approach for structural problems is the lower approximation order of derived properties like the strains and stresses (cf. chapter 2.6 in vol.1 of [137]).

Ways to circumvent the locking problem include the use of higher order elements or stabilisation techniques (cf. [9], vol.1 of [137], [112]). Mixed finite element methods (cf. [30]) provide another way to overcome these difficulties. Furthermore they can lead to better approximation properties in the values of interest like the stresses (cf. [30]). On the other hand mixed methods do not possess the stability of the standard Galerkin formulation anymore. Normally they require either *inf-sup* stable element pairs or again some stabilisation techniques (cf. [30]).

First order least squares finite element methods always lead to a minimisation problem and therefore to a symmetric positive definite system of equations for nonsingular problems. As the first order formulations usually include additional unknowns, which are related to the values of interest like the stresses, the approximation properties of these unknowns could be better than those for the standard Galerkin formulation. Furthermore some formulations were shown to give approximation properties independent of the Poisson ratio ν_s, which measures the incompressibility of the material. Drawbacks of the least squares methods are the higher number of unknowns and the higher regularity demands on the solution when used in the standard L^2 minimisation framework.

Within the context of this work, the most important feature of the least squares formulations for structural problems is the possibility to couple these formulations with the least squares formulations for the fluid part (cf. chapter 6) by some additional least squares terms representing the coupling conditions.

5.2 Current State of Research

The literature about first order LSFEM for the equations of linear elasticity can currently be divided into three main directions. As the equations of linear elasticity are equivalent to the compressible Stokes equations, all methods for the Stokes equations could theoretically be used for the equations of linear elasticity. These methods will be listed in the first section. After that a comparatively new approach which uses the displacement gradient as unknown will be shown. Finally a short overview about a recent approach using the real stress tensor as unknown will be given.

5.2.1 Formulations Based on the Stokes Equations

A possible first order formulation for the Stokes equations is the $up\omega$-formulation . But one problem with this formulation is the need for mesh dependent weights. While this is tolerable for CFD-applications, where the material parameters usually are the same in the complete domain, it is a real disadvantage in structural mechanics, where also different material parameters could appear in different regions of the domain. A method based on a functional, which avoids the mesh dependent weights by the use of a discrete

negative norm was examined in [34]:

$$\begin{aligned}\mathcal{J}(\boldsymbol{u},\omega,p) = \frac{1}{2}(\|f - (\nu\nabla\times\omega + (1+\nu\delta)\nabla\tilde{p}\|^2_{-1,\Omega} + \\ \nu^2\|\nabla\times\boldsymbol{u} - \omega\|^2_{0,\Omega} + \nu^2\|\nabla\cdot\boldsymbol{u} + \delta\tilde{p}\|^2_{0,\Omega}).\end{aligned} \tag{5.1}$$

In this formulation $\nu = \frac{\mu_s}{\lambda+\mu_s}$ and δ are used to switch between the incompressible Stokes equations where $\delta = 0$ and the equations of linear elasticity with $\delta = 1$. Functional Eq. (5.1) can be efficiently preconditioned by standard elliptic preconditioners and hence leads to an efficient numerical method (cf. [34]).

Two other Stokes formulations, which lead to optimal accuracy without mesh dependent weights are the augmented uUp-formulation (Eq. (4.25)), and the same formulation utilising a negative norm approach (Eq. (4.21)). Their application to linear elasticity is examined in [35]. Due to the coercivity results, which are independent of the Lamé constants, efficient preconditioners are available for this formulation as well.

Another least squares formulation, which is based on a Stokes like formulation for the equations of linear elasticity, is the negative norm approach shown in [25]. Here the minimised functional includes the second derivatives, which are still in H^{-1} if normal C_0 finite elements are considered. Without going into detail, the essential functional is:

$$\begin{aligned}\mathcal{J}(\boldsymbol{u},p) = \frac{1}{2}(\|\mathcal{L}(\boldsymbol{u},p)\|_{-1,\Omega} + \mu_0\|\nabla\cdot\boldsymbol{u} + \gamma p\|_{0,\Omega} + \|\sigma_n\|_{h,\Gamma_N} + \\ \|[\sigma_n]\|_{h,I} + \|\mathcal{K}(\boldsymbol{u},p)\|_h)\end{aligned} \tag{5.2}$$

where $\|\cdot\|_{h,\Gamma_N}$ and $\|\cdot\|_{h,I}$ denote special norms, which are defined on the element edges of the triangulation. $[\sigma_n]$ denotes the jump term of the stresses across the element boundary, $\mathcal{L}$ is the differential operator associated with the divergence of the stress tensor and $\mathcal{K}$ belongs to $\gamma p + \nabla\cdot\boldsymbol{u} = 0$, which is measured in a L^2 norm, weighted by the characteristic element size h. The parameter μ_0 provides a bound for the Lamé parameter μ_s, hence $0 < \mu_0 \leq \mu_s(x) \leq c_1\mu_0$ for an arbitrary $c_1 > 0$

For this functional Bramble et.al. prove the following estimate (cf. [25]):

$$c_2(\mu_0\|\boldsymbol{u}\|_{1,\Omega} + \|p\|_{0,\Omega}) \leq \mathcal{J}(\boldsymbol{u},p) \tag{5.3}$$

for some $c_2 > 0$ independent of h. A similar upper bound is easy to find and leads to an efficient elliptic preconditioner which can be used to construct a numerical scheme with an overall complexity of $O(n)$, where n is the number of unknowns. Numerical tests confirm the results. An interesting feature of this method is that in contrast to the other least squares methods, discontinuous piecewise constant functions could be used as approximation space for the pressure.

5.2.2 Displacement, Displacement Gradient Formulation

It is not clear, whether the first order formulation, which introduces the displacement gradient $\mathbf{H} = \nabla\boldsymbol{u}$ as a new unknown, was first proposed by Yang et.al. (cf. [135]) or by

Cai et.al. (cf. [44] and [88]) as those articles do not refer to each other and appeared approximately at the same time. The equations of linear elasticity together with mixed boundary conditions then become (cf. [43]):

$$\begin{aligned} \mathbf{H} - \nabla \boldsymbol{u} &= 0 \quad \text{in} \quad \Omega && (5.4)\\ -\nabla \cdot \mathbb{A}\mathbf{H} &= \boldsymbol{f} \quad \text{in} \quad \Omega && (5.5)\\ \nabla \times \mathbf{H} &= 0 \quad \text{in} \quad \Omega && (5.6)\\ \boldsymbol{n} \cdot \mathbb{A}\mathbf{H} &= 0 \quad \text{on} \quad \Gamma_N && (5.7)\\ \boldsymbol{t} \cdot \mathbf{H} &= 0 \quad \text{on} \quad \Gamma_D && (5.8) \end{aligned}$$

where $\boldsymbol{t}$ denotes the counterclockwise oriented unit tangent vector on the boundary. This basic formulation leads to different numerical schemes. Because $\boldsymbol{u}$ appears only in Eq. (5.4) it is possible to remove it from the system of equations and recover it later by solving only Eq. (5.4). This *two-stage* approach can be implemented with two different functionals for Eqs. (5.5-5.6). The first one uses negative norms and was proposed in [44]:

$$\mathcal{J}_{-1}(\mathbf{H}) = \|\boldsymbol{f} + \nabla \cdot \mathbb{A}\mathbf{H}\|^2_{-1^*,\Omega} + \|\nabla \times \mathbf{H}\|^2_{-1,\Omega}. \tag{5.9}$$

where $\|\cdot\|_{-1^*,\Omega}$ denotes the norm of the dual space of $H^1(\Omega)$. Alternatively the following functional based on L^2 norms can be used (cf.[135] and [44]):

$$\mathcal{J}_0(\mathbf{H}) = \|\boldsymbol{f} + \nabla \cdot \mathbb{A}\mathbf{H}\|^2_{0,\Omega} + \|\nabla \times \mathbf{H}\|^2_{0,\Omega}. \tag{5.10}$$

The last functional, which was proposed in [44] solves for $\boldsymbol{u}$ and $\mathbf{H}$ simultaneously:

$$\mathcal{J}(\mathbf{H}, \boldsymbol{u}) = \|\boldsymbol{f} + \nabla \cdot \mathbb{A}\mathbf{H}\|^2_{0,\Omega} + \|\nabla \times \mathbf{H}\|^2_{0,\Omega} + (\mathbb{A}(\mathbf{H} - \nabla \boldsymbol{u}), \mathbf{H} - \nabla \boldsymbol{u})_{0,\Omega} \tag{5.11}$$

These three functionals were extended to the 3D case in [88].

In [135] using the theory developed by Wendland for the LSFEM (cf. [129]), the following estimates are presented for the 2D case:

$$\|\mathbf{H}\|_{q+1,\Omega} \leq c_1(\|\mathcal{L}\mathbf{H}\|_{q,\Omega} + \|\mathcal{R}\mathbf{H}\|_{q+1/2,\Gamma}) \tag{5.12}$$

where $q \geq 0$ and $\mathcal{L}$ denotes Eqs. 5.5-5.6 and $\mathcal{R}$ denotes the boundary terms Eqs. 5.7-5.8. Clearly $\mathbf{H}$ must have sufficient regularity for this estimate to hold, i.e. $\mathbf{H} \in H^{q+1}(\Omega)^{d^2}$ with d being the spatial dimension. In the other parts of that article this estimate provides the basis for showing optimal convergence rates of the stresses with standard finite elements. These results are confirmed by numerical examples.

Slightly different estimates are derived in [44]. The boundary conditions are imposed on the solution spaces:

$$c_2(\|\mathbf{H}\|^2_{1,\Omega} + \lambda^2 \|\nabla \operatorname{tr} \mathbf{H}\|^2_{0,\Omega}) \leq \mathcal{J}_0(\mathbf{H}) \leq c_3(\|\mathbf{H}\|^2_{1,\Omega} + \lambda^2 \|\nabla \operatorname{tr} \mathbf{H}\|^2_{0,\Omega}). \tag{5.13}$$

For functional $\mathcal{J}_{-1}$ we have:

$$c_4(\|\mathbf{H}\|_{0,\Omega}^2 + \lambda^2\|\nabla \operatorname{tr}\mathbf{H}\|_{0,\Omega}^2) \leq \mathcal{J}_{-1}(\mathbf{H}) \leq c_5(\|\mathbf{H}\|_{0,\Omega}^2 + \lambda^2\|\nabla \operatorname{tr}\mathbf{H}\|_{0,\Omega}^2). \tag{5.14}$$

Finally for $\mathcal{J}$:

$$\begin{aligned} c_6(\|\mathbf{H}\|_{1,\Omega}^2 + \lambda^2\|\nabla \operatorname{tr}\mathbf{H}\|_{0,\Omega}^2 + \lambda\|\operatorname{tr}\mathbf{H}\|_{0,\Omega}^2 + \|\mathbb{A}^{1/2}\nabla\boldsymbol{u}\|_{0,\Omega}^2) \\ \leq \mathcal{J}(\mathbf{H},\boldsymbol{u}) \leq c_7(\|\mathbf{H}\|_{1,\Omega}^2 + \lambda^2\|\nabla \operatorname{tr}\mathbf{H}\|_{0,\Omega}^2 + \lambda\|\operatorname{tr}\mathbf{H}\|_{0,\Omega}^2 + \|\mathbb{A}^{1/2}\nabla\boldsymbol{u}\|_{0,\Omega}^2). \end{aligned} \tag{5.15}$$

Again sufficient regularity of $\mathbf{H}$ and $\boldsymbol{u}$ is required for these estimates.

The efficient solution of these equations with standard multigrid algorithms depends on the weak coupling between the different parts of the equations (cf. explanations in [67] for the advection diffusion equation). But in the estimates shown above, the term $\operatorname{tr}\mathbf{H}$ depends on λ and hence becomes dominant as $\lambda \to \infty$ (this happens when the material parameters approach incompressibility). As $\operatorname{tr}\mathbf{H} = H_{11} + H_{22}$ (with $\mathbf{H} = (H_{11}, H_{12}, H_{21}, H_{22})^T$) this implies a strong coupling between these unknowns and hence will destroy the diagonal dominance of the system. To overcome this problem a rotation $\mathbb{Q}$ is applied to $\mathbf{H}$, which does not impose major problems as the displacements have to be recovered in the second step anyway.

For the formulation using the rotated set of unknowns $\mathbf{V} = (V_{11}, V_{12}, V_{21}, V_{22})^T$, the following estimate holds:

$$c_8(\|\mathbf{V}\|_{1,\Omega}^2 + \lambda^2\|\nabla V_{11}\|_{0,\Omega} \leq \mathcal{J}_0(\mathbb{Q}\mathbf{V}) \leq c_9(\|\mathbf{V}\|_{1,\Omega}^2 + \lambda^2\|\nabla V_{11}\|_{0,\Omega} \tag{5.16}$$

Here the unknowns are clearly better decoupled, as only the gradients of V_{11} appear in the L^2-norm.

Numerical results for these formulations were published in [43] and in [135]. Both publications demonstrated that this method is able to handle even nearly incompressible material ($\nu_s \approx 0.5$) without convergence problems. In [43] also the performance of various multigrid schemes was examined. The authors were able to numerically confirm convergence rates of the multigrid scheme which were independent of h and the Lamé parameter λ.

5.2.3 Displacement-Stress Formulation

Looking at the basis of the equations of linear elasticity, they stem from the following equation:

$$\nabla \cdot \sigma = -\boldsymbol{f} \quad \text{in} \quad \Omega. \tag{5.17}$$

In this equation σ denotes the Cauchy stress tensor. Inserting the following constitutive equation, which is a linear relation between the stresses σ and the linearised Green strain tensor $\varepsilon(\boldsymbol{u})$:

$$\sigma = \mathbb{C}\varepsilon(\boldsymbol{u}) \tag{5.18}$$

into Eq. (5.17) gives the equations of linear elasticity (cf. section 2.4.2). $\mathbb{C}$ denotes the elasticity tensor. Closely related is the compliance tensor $\mathbb{E}$, which is the inverse of $\mathbb{C}$:

$$\varepsilon(\boldsymbol{u}) = \mathbb{E}\sigma \tag{5.19}$$

As Eq. (5.17) is already first order, it is a natural choice for a first order least squares formulation. The complete set of equations is then (cf. [40]):

$$\begin{aligned} \sigma - \mathbb{C}\varepsilon(\boldsymbol{u}) &= 0 \quad \text{in} \quad \Omega, & (5.20)\\ \nabla \cdot \sigma &= \boldsymbol{f} \quad \text{in} \quad \Omega. & (5.21) \end{aligned}$$

Alternatively Eq. (5.19) can be included and leads to (cf. [38]):

$$\begin{aligned} \mathbb{E}\sigma - \varepsilon(\boldsymbol{u}) &= 0 \quad \text{in} \quad \Omega, & (5.22)\\ \nabla \cdot \sigma &= \boldsymbol{f} \quad \text{in} \quad \Omega. & (5.23) \end{aligned}$$

As the elasticity tensor blows up for nearly incompressible material ($\nu_s \to 0.5$) the latter formulation is preferable in that case.

The following functionals were proposed for the use in a least squares finite element method from these equations (cf. [40], [38], [37]):

$$\mathcal{J}_1(\sigma, \boldsymbol{u}) = \|\nabla \cdot \sigma + \boldsymbol{f}\|^2_{0,\Omega} + \mu_s \|\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{u})\|^2_{0,\Omega}, \tag{5.24}$$

$$\mathcal{J}_2(\sigma, \boldsymbol{u}) = \|\nabla \cdot \sigma + \boldsymbol{f}\|^2_{0,\Omega} + \|\mathbb{E}\sigma - \varepsilon(\boldsymbol{u})\|^2_{0,\Omega}, \tag{5.25}$$

and

$$\mathcal{J}_{-1}(\sigma, \boldsymbol{u}) = \|\nabla \cdot \sigma + \boldsymbol{f}\|^2_{-1,\Omega} + \|\mathbb{E}\sigma - \varepsilon(\boldsymbol{u})\|^2_{0,\Omega}. \tag{5.26}$$

Actually the first two functionals $\mathcal{J}_1$ and $\mathcal{J}_2$ are very similar and correspond to the two different ways of formulating the relation between stresses and strains. Functional $\mathcal{J}_{-1}$ uses a negative norm to reduce the regularity demands on the right hands side from $\boldsymbol{f} \in L^2(\Omega)$ to $\boldsymbol{f} \in H^{-1}(\Omega)$, which is more natural in the context of linear elasticity as is can be shown that the weak form of the equation possesses a solution $\boldsymbol{u} \in H^1(\Omega)$ for $\boldsymbol{f} \in H^{-1}(\Omega)$ (cf. [51]).

In the following the two subspaces of $H^1(\Omega)^2$ and $H(\mathrm{div};\Omega)^2$ are required:

$$\begin{aligned} H^1_D(\Omega)^2 &= \{\boldsymbol{v} \in H^1(\Omega)^2 | \boldsymbol{v} = 0 \quad \forall \boldsymbol{x} \in \Gamma_D\} & (5.27)\\ H_N(\mathrm{div};\Omega)^2 &= \{\tau \in H(\mathrm{div};\Omega)^2 | \boldsymbol{n} \cdot \tau = 0 \quad \forall \boldsymbol{x} \in \Gamma_N\} & (5.28) \end{aligned}$$

Later also the dual space of $H^1_D(\Omega)$, $(H^1_D(\Omega))^*$ will be needed. Its norm is denoted by $\|\cdot\|_{-1,D,\Omega}$ (cf. Eq. (3.13)). For the bilinear form stemming from $\mathcal{J}_1$, coercivity can then be shown in the product space $H_N(\mathrm{div};\Omega)^2 \times H^1_D(\Omega)^2$ with the following scaled norm (cf. [40]):

$$\|(\tau, \boldsymbol{v})\|_{M_1,\Omega} = \left(\|\nabla \cdot \tau\|^2_{0,\Omega} + \mu_s \|\mathbb{C}^{-1/2}\tau\|^2_{0,\Omega} + \mu_s \|\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v})\|^2_{0,\Omega}\right)^{1/2}. \tag{5.29}$$

For $\mathcal{J}_2$ the norm is:

$$\|(\tau, \boldsymbol{v})\|_{M_2,\Omega} = \|\varepsilon(\boldsymbol{v})\|^2_{0,\Omega} + \|\tau\|^2_{0,\Omega} + \|\nabla \cdot \tau\|^2_{0,\Omega}, \tag{5.30}$$

and for $\mathcal{J}_{-1}$:

$$\|(\tau, \boldsymbol{v})\|_{M_{-1},\Omega} = \|\varepsilon(\boldsymbol{v})\|^2_{0,\Omega} + \|\tau\|^2_{0,\Omega}. \tag{5.31}$$

Using these norms, the following estimates can be established:

$$c_1\|(\tau, \boldsymbol{v})\|_{M_1,\Omega} \leq \quad \mathcal{J}_1(\tau, \boldsymbol{v}) \quad \leq c_2\|(\tau, \boldsymbol{v})\|_{M_1,\Omega} \tag{5.32}$$

$$c_3\|(\tau, \boldsymbol{v})\|_{M_2,\Omega} \leq \quad \mathcal{J}_2(\tau, \boldsymbol{v}) \quad \leq c_4\|(\tau, \boldsymbol{v})\|_{M_2,\Omega} \tag{5.33}$$

$$c_5\|(\tau, \boldsymbol{v})\|_{M_{-1},\Omega} \leq \quad \mathcal{J}_{-1}(\tau, \boldsymbol{v}) \quad \leq c_6\|(\tau, \boldsymbol{v})\|_{M_{-1},\Omega}. \tag{5.34}$$

Due to some specific properties (for details refer to [40]), the functional $\mathcal{J}_1$ needs special finite element approximations for the displacements in the incompressible limit. Therefore the use of nonconforming Crouzeix-Raviart elements or their quadratic counterpart, the elements from Fortin and Soulie (cf.[68]) are used to approximate the displacement. The slightly different functional $\mathcal{J}_2$ allows also the use of standard finite elements for the displacement, even in the incompressible limit. The stresses have to be approximated by finite subspaces of $H(\mathrm{div};\Omega)$ in both cases. Raviart-Thomas elements are used for this purpose.

Numerical tests in [40] confirm the result that the convergence rates deteriorate in the incompressible limit, if conforming standard finite elements are used in the incompressible limit with functional $\mathcal{J}_1$. In contrast the functional $\mathcal{J}_2$ used in [38] gives optimal convergence rates in the incompressible limit even together with the standard conforming finite elements.

Both articles use the least squares functional as built-in error estimator and construct an adaptive algorithm, which is based on that estimator.

5.3 Galerkin/Newmark Formulation

As structural problems were one of the first applications of the Galerkin FEM, these formulations are well understood and a vast amount of literature exists for this formulation. For some examples, a least squares fluid formulation is coupled with a Galerkin structure formulation. Hence a very short description of the used formulation will be presented in this section. A more detailed description can be found in section A.2.

The definition of the mass and stiffness matrices $\mathbf{M}$ and $\mathbf{K}$ is shown in the appendix and can also be found in most introductory textbooks (for example [5],[137]). For time integration the variant of the *Newmark* scheme [5] which corresponds to the trapezoidal rule is used, as it is stable and energy conserving for linear systems. As nonlinear structures are out of the scope of this work, the limitation of the Newmark method to linear systems is unproblematic.

Normally new accelerations are computed in each *Newmark* time-step by solving a system of equations. Velocities and displacements are then computed from the accelerations. Because velocities are required in the coupled system of equations, the scheme was rearranged to have the velocities as unknowns. For the sake of simplicity also the displacements were included into the global system of equations for the coupled problem. In the examples which will be shown in this work, the number of unknowns in the structure is negligible compared to the unknowns in the fluid. Thus using this approach does not have a significant impact on the computational efficiency and the correct displacements are directly available for grid deformation. This leads then to the following system of equations for the structural part:

$$\begin{pmatrix} c_1\mathbf{K}+\mathbf{M} & 0 \\ c_2\mathbf{K} & \mathbf{M} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{n+1} \\ \boldsymbol{v}_{n+1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{g}_{n+1} \\ \boldsymbol{h}_{n+1} \end{pmatrix} \tag{5.35}$$

with $c_1 = \frac{\Delta t^2 * (1/2-\beta)}{\rho_s}$, $c_2 = \frac{\Delta t * \gamma}{\rho_s}$ and

$$\boldsymbol{g_{n+1}} = \mathbf{M}\boldsymbol{u}_n + \Delta t \mathbf{M}\boldsymbol{v}_n - \beta \Delta t^2 \frac{1}{\rho_s}\mathbf{K}\boldsymbol{u}_n + \Delta t^2 \boldsymbol{f} \tag{5.36}$$

$$\boldsymbol{h_{n+1}} = \mathbf{M}\boldsymbol{v}_n + \Delta t(1-\gamma)\frac{1}{\rho_s}\mathbf{K}\boldsymbol{u}_n + \Delta t \boldsymbol{f} \tag{5.37}$$

where $\boldsymbol{f}$ denotes external or volume forces, and $\beta = 1/4$, $\gamma = 1/2$.

The model can cope with the plane strain and the plane stress assumption. Usually the plane strain assumption is better for a meaningful 2D-model. Nevertheless in some examples the plane stress assumption is used as well in compliance with the setting of the original test problem.

It is a well known fact that linear ansatz functions lead to elements, which are too "stiff" if no special techniques are applied (cf. [11]). Here this difficulty is overcome by using higher order elements, like isoparametric 9 or 16 node elements with quadratic ($Q2$) or cubic ($Q3$) shapefunctions, which proved to be quite accurate (cf. subsection 5.10.1).

5.4 First Order LSFEM for the Wave Equation

As the theory for the LSFEM applied to initial boundary value problems is not very well developed, the following sections will introduce the basic concepts which will later be used to analyse formulations for the equations describing the dynamic behaviour of linear elastic solids. For this purpose, the wave equation serves as an example. As it is a second order hyperbolic equation, it is similar to the equations of linear elastodynamics.

First the 1D wave equation will be analysed to demonstrate the principal path of showing stability. After that these results will be extended to the anisotropic wave equation in 2D and 3D. The results presented for that equation demonstrate, that the

least squares FEM is not able to guarantee stability in the H^1 norms, which is required for the standard finite elements. Instead a stability result in the weaker $H(\mathrm{div})$ norm can be found. To get stability estimates in the stronger H^1 norm, the equation must be augmented by an additional term. This augmented or stabilised method is analysed in the third part of this section.

5.4.1 The 1D Wave Equation

In this section the wave equation in 1D will serve as an introductory example to demonstrate the basic concepts and the problems which occur when the first order LSFEM is used to solve second order hyperbolic problems. It reads:

$$\begin{aligned} \frac{\partial u^2}{\partial t^2} - c^2 \frac{\partial u^2}{\partial x^2} &= f \quad \text{in} \quad \Omega && (5.38)\\ u &= 0 \quad \text{on} \quad \Gamma_1 && (5.39)\\ \frac{\partial u}{\partial x} &= 0 \quad \text{on} \quad \Gamma_2 && (5.40)\\ u(0,x) &= u_0(x) \quad \text{on} \quad \Omega. && (5.41) \end{aligned}$$

with $\Omega = [0 \ldots 1]$ and $\Gamma_1 = 0, \Gamma_2 = 1$. The function $u_0(x)$ denotes the initial conditions. For simplicity the boundary conditions are assumed to be constant in time and the volume force is set to $f = 0$.

To simplify the notation some definitions will be introduced. These definitions refer to Eq. (3.97). Using the class of θ-methods, the following differential operators play a major role:

$$\mathcal{E}_l = \mathcal{I} + \Delta t \theta \mathcal{L}, \quad \mathcal{E}_r = \mathcal{I} - \Delta t (1-\theta) \mathcal{L} \tag{5.42}$$

With these definitions, the general variational statement of the L^2 LSFEM for one time step reads:

$$(\mathcal{E}_l \boldsymbol{u}_{n+1}, \mathcal{E}_l \boldsymbol{v})_{0,\Omega} = (\mathcal{E}_r \boldsymbol{u}_n, \mathcal{E}_l \boldsymbol{v})_{0,\Omega} + (\theta \boldsymbol{f}_{n+1} + (1-\theta)\boldsymbol{f}_n, \mathcal{E}_l \boldsymbol{v})_{0,\Omega} \quad \forall \boldsymbol{v} \in \mathbf{V}(\Omega) \tag{5.43}$$

where the Hilbert space $\mathbf{V}(\Omega)$ depends on the problem. The following lemma will be useful for the energy estimates which will be used to analyse the temporal development of the solution found by the LSFEM.

Lemma 3 *If a solution* $\boldsymbol{u_{n+1}} \in \mathbf{V}(\Omega)$ *exists and satisfies Eq. (5.43), then*

$$\|\mathcal{E}_l \boldsymbol{u}_{n+1}\|_{0,\Omega} \leq \|\mathcal{E}_r \boldsymbol{u}_n\|_{0,\Omega} + \|\theta \boldsymbol{f}_{n+1} + (1-\theta)\boldsymbol{f}_n\|_{0,\Omega} \tag{5.44}$$

Proof: As Eq. (5.43) holds for all $\boldsymbol{v} \in \mathbf{V}(\Omega)$, and $\boldsymbol{u}_{n+1} \in \mathbf{V}(\Omega)$, it holds also for $\boldsymbol{v} = \boldsymbol{u}_{n+1}$. Hence:

$$\|\mathcal{E}_l \boldsymbol{u}_{n+1}\|^2_{0,\Omega} = (\mathcal{E}_r \boldsymbol{u}_n + \theta \boldsymbol{f}_{n+1} + (1-\theta)\boldsymbol{f}_n, \mathcal{E}_l \boldsymbol{u}_{n+1})_{0,\Omega}. \tag{5.45}$$

Using the Cauchy-Schwarz inequality:

$$\|\mathcal{E}_l \boldsymbol{u}_{n+1}\|_{0,\Omega}^2 \leq \|\mathcal{E}_r \boldsymbol{u}_n + \theta \boldsymbol{f}_{n+1} + (1-\theta)\boldsymbol{f}_n\|_{0,\Omega} \cdot \|\mathcal{E}_l \boldsymbol{u}_{n+1}\|_{0,\Omega}, \tag{5.46}$$

and dividing by $\|\mathcal{E}_l \boldsymbol{u}_{n+1}\|_{0,\Omega}$:

$$\|\mathcal{E}_l \boldsymbol{u}_{n+1}\|_{0,\Omega} \leq \|\mathcal{E}_r \boldsymbol{u}_n + \theta \boldsymbol{f}_{n+1} + (1-\theta)\boldsymbol{f}_n\|_{0,\Omega}, \tag{5.47}$$

which in conjunction with the triangle inequality shows the result. □

To transform the second order wave equation Eq. (5.38) into an equivalent first order formulation, the temporal and spatial derivatives of u will be introduced as new unknowns $v = \partial u/\partial t$ and $p = c\partial u/\partial x$:

$$\frac{\partial v}{\partial t} - c\frac{\partial p}{\partial x} = 0 \quad \text{in} \quad \Omega \tag{5.48}$$

$$\frac{\partial p}{\partial t} - c\frac{\partial v}{\partial x} = 0 \quad \text{in} \quad \Omega \tag{5.49}$$

Applying the standard temporal finite difference discretisation, the semidiscrete form of Eq. (5.48) becomes:

$$v_{n+1} - k_1 c\frac{\partial p_{n+1}}{\partial x} = v_n + k_2 c\frac{\partial p_n}{\partial x} \quad \text{in} \quad \Omega \tag{5.50}$$

$$p_{n+1} - k_1 c\frac{\partial v_{n+1}}{\partial x} = p_n + k_2 c\frac{\partial v_n}{\partial x} \quad \text{in} \quad \Omega \tag{5.51}$$

with $k_1 = \theta\Delta t$ and $k_2 = (1-\theta)\Delta t$. For this first order system the following least squares functional can be defined. It is the basis for the problem which has to be solved in each time step:

$$\mathcal{J}(v_{n+1}, p_{n+1}) = \frac{1}{2}\left(\left\|v_{n+1} - k_1 c\frac{\partial p_{n+1}}{\partial x} - f_1\right\|_{0,\Omega}^2 + \left\|p_{n+1} - k_1 c\frac{\partial v_{n+1}}{\partial x} - f_2\right\|_{0,\Omega}^2\right) \tag{5.52}$$

with $f_1 = v_n + k_2 c\frac{\partial p_n}{\partial x}$ and $f_2 = p_n + k_2 c\frac{\partial v_n}{\partial x}$. Finally the following bilinear form (dropping the index n) can be found:

$$\begin{aligned}\mathcal{B}((v,p),(v^1,v^2)) = &\left(v - k_1 c\frac{\partial p}{\partial x}, v^1 - k_1 c\frac{\partial v^2}{\partial x}\right)_{0,\Omega} + \\ &\left(p - k_1 c\frac{\partial v}{\partial x}, v^2 - k_1 c\frac{\partial v^1}{\partial x}\right)_{0,\Omega}.\end{aligned} \tag{5.53}$$

Stability guarantees that the solution at an arbitrary time instance is bounded in some norm by terms including the initial conditions, the boundary conditions and the right hand side term f. The following theorem shows this property for the least squares formulation based on Eq. (5.52).

Theorem 4 *Assuming $f = 0$ and $\theta = 0.5$, the following estimate holds for the solution v_{N+1}, p_{N+1}, found by applying N times the variational statement based on Eq. (5.52) with initial conditions $(v_0, p_0) \in H^1(\Omega) \times H^1(\Omega)$:*

$$\begin{aligned}\|v_{N+1}\|^2_{0,\Omega}+\|p_{N+1}\|^2_{0,\Omega} + (k_1c)^2\|\frac{\partial v_{N+1}}{\partial x}\|^2_0 + (k_1c)^2\|\frac{\partial p_{N+1}}{\partial x}\|^2_0 \leq \\ \|v_0\|^2_{0,\Omega} + \|p_0\|^2_{0,\Omega} + (k_1c)^2\|\frac{\partial v_0}{\partial x}\|^2_{0,\Omega} + (k_1c)^2\|\frac{\partial p_0}{\partial x}\|^2_{0,\Omega}+ \\ 2(k_1c)\sum_{n=0}^{N}([v_np_n]^b_a + [v_{n+1}p_{n+1}]^b_a).\end{aligned} \tag{5.54}$$

Here $[f]^b_a$ denotes the difference $f(b) - f(a)$ for a function f.

Proof: First applying lemma 3 to the bilinear forms stemming from 5.52 the following estimate can be found:

$$\|v_{n+1} - k_1c\frac{\partial p_{n+1}}{\partial x}\|^2_{0,\Omega} + \|p_{n+1} - k_1c\frac{\partial v_{n+1}}{\partial x}\|^2_{0,\Omega} \leq \|v_n + k_2c\frac{\partial p_n}{\partial x}\|^2_{0,\Omega} + \|p_n + k_2c\frac{\partial v_n}{\partial x}\|^2_{0,\Omega}. \tag{5.55}$$

To get the desired estimate the left hand term of Eq. (5.55) must be greater than some norm of (v, p), while the right hand side term has to be less than some norm of (v, p). First the lower bound will be shown. The left hand side of Eq. (5.55) is equivalent to:

$$\begin{aligned}\|v_{n+1}\|^2_{0,\Omega} + \|k_1c\frac{\partial p_{n+1}}{\partial x}\|^2_{0,\Omega} - 2k_1c(v_{n+1}, \frac{\partial p_{n+1}}{\partial x})_{0,\Omega}+ \\ \|p_{n+1}\|^2_{0,\Omega} + \|k_1c\frac{\partial v_{n+1}}{\partial x}\|^2_{0,\Omega} - 2k_1c(p_{n+1}, \frac{\partial v_{n+1}}{\partial x})_{0,\Omega}.\end{aligned} \tag{5.56}$$

Using integration by parts it becomes:

$$\|v_{n+1}\|^2_{0,\Omega} + \|k_1c\frac{\partial p_{n+1}}{\partial x}\|^2_{0,\Omega} + \|p_{n+1}\|^2_{0,\Omega} + \|k_1c\frac{\partial v_{n+1}}{\partial x}\|^2_{0,\Omega} - 2k_1c[v_{n+1}p_{n+1}]^b_a. \tag{5.57}$$

Using a similar procedure for the right hand side, the following expression for the right hand side term in Eq. (5.55) can be found:

$$\|v_n\|^2_{0,\Omega} + \|k_2c\frac{\partial p_n}{\partial x}\|^2_{0,\Omega} + \|p_n\|^2_{0,\Omega} + \|k_2c\frac{\partial v_n}{\partial x}\|^2_{0,\Omega} + 2k_2c[v_np_n]^b_a. \tag{5.58}$$

Combing Eq. (5.57) and Eq. (5.58):

$$\begin{aligned}\|v_{n+1}\|^2_{0,\Omega} + \|k_1c\frac{\partial p_{n+1}}{\partial x}\|^2_{0,\Omega} + \|p_{n+1}\|^2_{0,\Omega} + \|k_1c\frac{\partial v_{n+1}}{\partial x}\|^2_{0,\Omega} \leq \\ \|v_n\|^2_{0,\Omega} + \|k_2c\frac{\partial p_n}{\partial x}\|^2_{0,\Omega} + \|p_n\|^2_{0,\Omega} + \|k_2c\frac{\partial v_n}{\partial x}\|^2_{0,\Omega} + 2k_1c[v_np_n]^b_a + 2k_2c[v_{n+1}p_{n+1}]^b_a.\end{aligned} \tag{5.59}$$

Summing both sides from $n = 0$ to $n = N$ together with $k_2 = k_1$ for $\theta = 0.5$ shows the theorem. □

For finite Δt Theorem (4) guarantees stability in the H^1 norm, which is required for the spatial finite element discretisation using normal elements. In the limit $\Delta t \to 0$ the derivatives of v and p will vanish. But as the L^2 stability of v and p is guaranteed, stability in the H^1-norm for u is conserved.

5.4.2 The 2- and 3D Wave Equation

While the 1D Wave equation could be treated with the least squares FEM without difficulties, the analysis of least squares methods for the same equation in 2- and 3D will reveal some interesting details and help to demonstrate some fundamental least squares techniques, which will later be applied to the equations of linear elastodynamics. The problem considered in this section is:

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \nabla \cdot c^2 \nabla u &= 0 \quad \text{in} \quad \Omega && (5.60)\\ u &= 0 \quad \text{on} \quad \Gamma_D && (5.61)\\ \frac{\partial u}{\partial \boldsymbol{n}} &= 0 \quad \text{on} \quad \Gamma_N && (5.62)\\ u(0) &= u_0. && (5.63) \end{aligned}$$

The domain is denoted by Ω with boundarys $\partial\Omega = \Gamma_N \cup \Gamma_D$, $\Gamma_D \cap \Gamma_N = \emptyset$, $\Gamma_D \neq \emptyset$. As before c denotes the wave propagation speed. This coefficient should be constant in the whole domain Ω.

An equivalent first order system is:

$$\begin{aligned} \frac{\partial \boldsymbol{p}}{\partial t} - c\nabla v &= 0 \quad \text{in} \quad \Omega && (5.64)\\ \frac{\partial v}{\partial t} - \nabla \cdot c\boldsymbol{p} &= 0 \quad \text{in} \quad \Omega && (5.65) \end{aligned}$$

where $v = \partial u/\partial t$ and $\boldsymbol{p} = c\nabla v$. Using the class of θ-methods for time discretisation, the semidiscrete system becomes:

$$\begin{aligned} \boldsymbol{p}_{n+1} - k_1 c\nabla v_{n+1} &= \boldsymbol{p}_n + k_2 c\nabla v_n \quad \text{in} \quad \Omega && (5.66)\\ v_{n+1} - k_1 \nabla \cdot c\boldsymbol{p}_{n+1} &= v_n + k_2 \nabla \cdot c\boldsymbol{p}_n \quad \text{in} \quad \Omega && (5.67) \end{aligned}$$

with $k_1 = \theta\Delta t$ and $k_2 = (1-\theta)\Delta t$.

The corresponding L^2 least squares functional is then

$$\mathcal{J}(\boldsymbol{p}_{n+1}, \boldsymbol{u}_{n+1}) = \frac{1}{2}(\|\boldsymbol{p}_{n+1} - k_1 c\nabla v_{n+1} - \boldsymbol{f}_1\|^2_{0,\Omega} + \|v_{n+1} - k_1 \nabla \cdot c\boldsymbol{p}_{n+1} - f_2\|^2_{0,\Omega}) \quad (5.68)$$

with $\boldsymbol{f}_1 = \boldsymbol{p}_n + k_2 c\nabla v_n$ and $f_2 = v_n + k_2 \nabla \cdot c\boldsymbol{p}_n$.

For the least squares formulation which is based on this functional, the following stability estimate holds:

Theorem 5 *Assuming $\theta = 0.5$, the following estimate holds for the solution $\boldsymbol{p}_{N+1}, v_{N+1}$, found by applying N times the variational statement based on Eq. (5.68) with initial conditions $\boldsymbol{p}_0 \in H(\mathrm{div};\Omega), v_0 \in H^1(\Omega)$:*

$$\begin{aligned}\|\boldsymbol{p}_{N+1}\|^2_{0,\Omega} + k_1^2\|\nabla \cdot c\boldsymbol{p}_{N+1}\|^2_{0,\Omega} + \|v_{N+1}\|^2_{0,\Omega} + k_1^2\|c\nabla v_{N+1}\|^2_{0,\Omega} \leq \\ \|\boldsymbol{p}_0\|^2_{0,\Omega} + k_1^2\|\nabla \cdot c\boldsymbol{p}_0\|^2_{0,\Omega} + \|v_0\|^2_{0,\Omega} + k_1^2\|c\nabla v_0\|^2_{0,\Omega} + \\ 2k_1 \sum_{n=0}^{N} \int_\Gamma (\boldsymbol{n} \cdot c\boldsymbol{p}_{n+1}) v_{n+1} + (\boldsymbol{n} \cdot c\boldsymbol{p}_n) v_n \, d\Gamma\end{aligned} \tag{5.69}$$

Proof: Applying lemma 3 to the bilinear forms stemming from Eq. (5.68) leads to the following estimate:

$$\begin{aligned}\|\boldsymbol{p}_{n+1} - k_1 c\nabla v_{n+1}\|^2_{0,\Omega} + \|v_{n+1} - k_1 \nabla \cdot c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} \leq \\ \|\boldsymbol{p}_n + k_2 c\nabla v_n\|^2_{0,\Omega} + \|v_n + k_2 \nabla \cdot c\boldsymbol{p}_n\|^2_{0,\Omega}\end{aligned} \tag{5.70}$$

First the left hand will be bounded from below. The left hand side term of Eq. (5.70) is

$$\begin{aligned}\|\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + k_1^2\|c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + \|v_{n+1}\|^2_{0,\Omega} + k_1^2\|c\nabla v_{n+1}\|^2_{0,\Omega} \\ - 2k_1(\boldsymbol{p}_{n+1}, c\nabla v_{n+1})_{0,\Omega} - 2k_1(v_{n+1}, \nabla \cdot c\boldsymbol{p}_{n+1})_{0,\Omega}\end{aligned} \tag{5.71}$$

The application of Green's theorem yields:

$$\begin{aligned}-2k_1(\boldsymbol{p}_{n+1}, c\nabla v_{n+1})_0 = -2k_1(c\boldsymbol{p}_{n+1}, \nabla v_{n+1})_0 = \\ 2k_1(\nabla \cdot c\boldsymbol{p}_{n+1}, v_{n+1})_{0,\Omega} - 2k_1\langle \boldsymbol{n} \cdot c\boldsymbol{p}_{n+1}, v_{n+1}\rangle_\Gamma\end{aligned} \tag{5.72}$$

which allows the simplification of the left hand side to:

$$\|\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + k_1^2\|c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + \|v_{n+1}\|^2_{0,\Omega} + k_1^2\|c\nabla v_{n+1}\|^2_{0,\Omega} - 2k_1\langle \boldsymbol{n} \cdot c\boldsymbol{p}_{n+1}, v_{n+1}\rangle_\Gamma \tag{5.73}$$

Using the same procedure for the right hand side of Eq. (5.70) gives:

$$\|\boldsymbol{p}_n\|^2_{0,\Omega} + k_2^2\|c\boldsymbol{p}_n\|^2_{0,\Omega} + \|v_n\|^2_{0,\Omega} + k_2^2\|c\nabla v_n\|^2_{0,\Omega} + 2k_2\langle \boldsymbol{n} \cdot c\boldsymbol{p}_n, v_n\rangle_\Gamma \tag{5.74}$$

Inserting these expressions for the left and right hand side into Eq. (5.70) and summing from $n = 0$ to $n = N$ then shows the theorem for $k_1 = k_2$. $\square$.

From the mathematical point of view, this theorem guarantees stability in the space $H^1(\Omega) \times H(\mathrm{div};\Omega)$ which is sufficient. Numerically the consequence is that the approximation spaces should be subspaces of $H^1(\Omega) \times H(\mathrm{div};\Omega)$. Otherwise the stability properties will be lost as the approximation functions have components which cannot be controlled by the least squares functional. Hence the gradient $\boldsymbol{p}$ should be approximated using Raviart-Thomas spaces.

5.4.3 An Augmented First Order Formulation for the 2- and 3D Wave Equation

To circumvent the use of Raviart-Thomas spaces, the functional Eq. (5.68) could be stabilised by adding $\|\nabla \times \boldsymbol{p}\|^2_{0,\Omega}$. A similar approach was suggested for the first order formulation of the Laplace equation (cf. subsection 3.5.3), but could be extended to the formulation for the wave equation in 2 and 3 dimensions.

First the equivalence of the augmented formulation and the original equation will be shown. Applying the curl operator $\nabla\times$ to Eq. (5.64), leads to the following equation:

$$\nabla \times \frac{\partial \boldsymbol{p}}{\partial t} - \nabla \times c\nabla v = \nabla \times \frac{\partial \boldsymbol{p}}{\partial t} = 0 \quad \text{in} \quad \Omega. \tag{5.75}$$

After temporal discretisation the following expression can be obtained:

$$\alpha\nabla \times \boldsymbol{p}_{n+1} = \alpha\nabla \times \boldsymbol{p}_n \quad \text{in} \quad \Omega \tag{5.76}$$

where the weight α can be chosen arbitrarily. As it will be shown later a properly selected weight is necessary for a sharp stability estimate. The method based on Eqs. (5.66–5.67) and Eq. (5.76) leads to the following functional, which is the basis for finding the solution of the next time step:

$$\begin{aligned}\mathcal{J}(\boldsymbol{p}_{n+1}, v_{n+1}) = \frac{1}{2}(&\|\boldsymbol{p}_{n+1} - k_1 c\nabla v_{n+1} - \boldsymbol{f}_1\|^2_{0,\Omega} + \|v_{n+1} - k_1\nabla \cdot c\boldsymbol{p}_{n+1} - f_2\|^2_{0,\Omega} + \\ &\|\alpha\nabla \times \boldsymbol{p}_{n+1} - f_3\|^2_{0,\Omega})\end{aligned} \tag{5.77}$$

with $f_3 = \alpha\nabla \times \boldsymbol{p}_n$ and $\boldsymbol{f}_1$ and f_2 being the same as for the unaugmented formulation.

A small lemma is required before the main part of the stability analysis (this is Lemma 5.1 in [83]):

Lemma 4 *Let Ω be a bounded and convex open subset of $\mathbb{R}^3$ with a boundary $\Gamma = \Gamma_1 \cup \Gamma_2$ and $\Gamma_1 \cap \Gamma_2 = \emptyset$. Then every function $\boldsymbol{u} \in (H^1(\Omega))^3$ with $\boldsymbol{n} \cdot \boldsymbol{u} = 0$ on Γ_1 and $\boldsymbol{n} \times \boldsymbol{u} = 0$ on Γ_2 satisfies:*

$$|\boldsymbol{u}|^2_1 \leq \|\nabla \cdot \boldsymbol{u}\|^2_{0,\Omega} + \|\nabla \times \boldsymbol{u}\|^2_{0,\Omega}. \tag{5.78}$$

Proof: A proof can be found in [83], p.83. □.

Now we come to the main theorem, which shows the stability of the augmented formulation for the wave equation in 2- and 3D.

Theorem 6 *With $\theta = 0.5$ and $\alpha = k_1 c$ the following estimate holds for the solution $\boldsymbol{p}_{N+1}, v_{N+1}$, found by applying N times the variational statement based on Eq. (5.68)*

with initial conditions $\boldsymbol{p}_0 \in H^1(\Omega)^d, v_0 \in H^1(\Omega)$*:*

$$\begin{aligned}\|\boldsymbol{p}_{N+1}\|^2_{0,\Omega}+k_1^2c^2|\boldsymbol{p}_{N+1}|^2_{1,\Omega} + \|v_{N+1}\|^2_{0,\Omega} + k_1^2c^2|v_{N+1}|^2_{1,\Omega} \leq \\ \|\boldsymbol{p}_0\|^2_{0,\Omega} + k_1^2c^2|\boldsymbol{p}_0|^2_{1,\Omega} + \|v_0\|^2_{0,\Omega} + k_1^2c^2|v_0|^2_{1,\Omega}+ \\ 2k_1\sum_{n=0}^{N}\int_\Gamma c(\boldsymbol{n}\cdot\boldsymbol{p}_{n+1})v_{n+1} + c(\boldsymbol{n}\cdot\boldsymbol{p}_n)v_n\,d\Gamma\end{aligned} \tag{5.79}$$

Proof: Using again lemma 3 for the bilinear forms stemming from functional Eq. (5.77) we find the basis for the analysis:

$$\begin{aligned}\|\boldsymbol{p}_{n+1} - k_1c\nabla v_{n+1}\|^2_{0,\Omega} + \|v_{n+1} - k_1\nabla\cdot c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + \|k_1\nabla\times c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} \leq \\ \|\boldsymbol{p}_n - k_2c\nabla v_n\|^2_{0,\Omega} + \|v_n - k_2\nabla\cdot c\boldsymbol{p}_n\|^2_{0,\Omega} + \|k_1\nabla\times c\boldsymbol{p}_n\|^2_{0,\Omega}\end{aligned} \tag{5.80}$$

Again the left hand side will be analysed first. Separating the terms inside the norms leads to:

$$\begin{aligned}\|\boldsymbol{p}_{n+1}\|^2_{0,\Omega}+\|k_1c\nabla v_{n+1}\|^2_{0,\Omega} + \|v_{n+1}\|^2_{0,\Omega} + \|k_1\nabla\cdot c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + \|k_1\nabla\times c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} \\ - 2k_1(\boldsymbol{p}_{n+1}, c\nabla v_{n+1})_{0,\Omega} - 2k_1(\boldsymbol{v}_{n+1}, \nabla\cdot c\boldsymbol{p}_{n+1})_{0,\Omega}.\end{aligned} \tag{5.81}$$

With Green's theorem the scalar products reduce to a boundary term:

$$\begin{aligned}\|\boldsymbol{p}_{n+1}\|^2_{0,\Omega}+\|k_1c\nabla v_{n+1}\|^2_{0,\Omega} + \|v_{n+1}\|^2_{0,\Omega} + \|k_1\nabla\cdot c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + \|k_1\nabla\times c\boldsymbol{p}_{n+1}\|^2_{0,\Omega} \\ - 2k_1\langle\boldsymbol{n}\cdot c\boldsymbol{p}_{n+1}, v\rangle_\Gamma.\end{aligned} \tag{5.82}$$

Now by using lemma 4 the equivalence to a H^1 like norm can be shown:

$$\|\boldsymbol{p}_{n+1}\|^2_{0,\Omega} + k_1^2\|\nabla c\boldsymbol{p}\|^2_{0,\Omega} + \|v_{n+1}\|^2_{0,\Omega} + k_1^2\|c\nabla v_{n+1}\|^2_{0,\Omega} - 2k_1\langle\boldsymbol{n}\cdot c\boldsymbol{p}_{n+1}, v\rangle_\Gamma. \tag{5.83}$$

Setting $k_2 = k_1$ and using the same procedure for the right hand side shows then the theorem. □

5.5 The *vm*-Formulation for Linear Elastodynamics

Interestingly, the stabilised wave equation is already very close to a first order formulation for the equations describing transient linear elastic problems. First the basic idea will be illustrated for the corresponding stationary problem. Hence in the following parts the displacements in x- and y-direction will be denoted by u_x, u_y. A vector unknown $\boldsymbol{m}$ is introduced:

$$\boldsymbol{m} = \mathbf{G}\nabla u_x + \mathbf{G}\nabla\times u_y \tag{5.84}$$

with

$$\mathbf{G} = \begin{pmatrix} \sqrt{(\lambda+2\mu_s)/\rho_s} & 0 \\ 0 & \sqrt{\mu_s/\rho_s}\end{pmatrix} \tag{5.85}$$

and the Lamé parameters λ and μ_s. The divergence of the linearised stress tensor, can now be expressed by $\boldsymbol{m}$.

$$\nabla \cdot \sigma = \begin{pmatrix} \nabla \cdot \mathbf{G}\boldsymbol{m} \\ -\nabla \times \mathbf{G}\boldsymbol{m} \end{pmatrix}. \tag{5.86}$$

Hence a first order formulation, which is equivalent to Eq. (2.45), reads:

$$\begin{aligned} \boldsymbol{m} - \mathbf{G}\nabla u_x - \mathbf{G}\nabla \times u_y &= 0 \quad \text{in} \quad \Omega && (5.87)\\ -\nabla \cdot \mathbf{G}\boldsymbol{m} &= 0 \quad \text{in} \quad \Omega && (5.88)\\ \nabla \times \mathbf{G}\boldsymbol{m} &= 0 \quad \text{in} \quad \Omega && (5.89) \end{aligned}$$

Now switching to the transient problem, $\boldsymbol{m}$ will be formulated in terms of the displacement rate or the velocity. The unknowns v_x, v_y will denote the velocity. Then the first order system, which is equivalent to Eq. (2.44) reads:

$$\begin{aligned} \frac{\partial \boldsymbol{m}}{\partial t} - \mathbf{G}\nabla v_x - \mathbf{G}\nabla \times v_y &= 0 \quad \text{in} \quad \Omega && (5.90)\\ \frac{\partial v_x}{\partial t} - \nabla \cdot \mathbf{G}\boldsymbol{m} &= 0 \quad \text{in} \quad \Omega && (5.91)\\ \frac{\partial v_y}{\partial t} + \nabla \times \mathbf{G}\boldsymbol{m} &= 0 \quad \text{in} \quad \Omega && (5.92) \end{aligned}$$

and should be accompanied by the following boundary conditions, which are to be defined on two non empty, non overlapping subsets Γ_1 and Γ_2 of the boundary Γ:

$$\begin{aligned} \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m} &= 0 \quad \text{on} \quad \Gamma_1 && (5.93)\\ v_y &= 0 \quad \text{on} \quad \Gamma_1 && (5.94)\\ \boldsymbol{n} \times \mathbf{G}\boldsymbol{m} &= 0 \quad \text{on} \quad \Gamma_2 && (5.95)\\ \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m} &= 0 \quad \text{on} \quad \Gamma_2 && (5.96) \end{aligned}$$

As the analysis will reveal, these boundary conditions are required for stability of the least squares formulation but will not correspond to a physical meaningful setup.

Discretising this formulation in time using the θ-methods, leads to

$$\begin{aligned} \boldsymbol{m}_{n+1} - k_1\mathbf{G}\nabla v_x - k_1\mathbf{G}\nabla \times v_y &= \boldsymbol{m}_n + k_2\mathbf{G}\nabla v_x + k_2\mathbf{G}\nabla \times v_y && (5.97)\\ v_x^{n+1} - k_1\nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1} &= v_x^n + k_2\nabla \cdot \mathbf{G}\boldsymbol{m}_n && (5.98)\\ v_y^{n+1} + k_1\nabla \times \mathbf{G}\boldsymbol{m}_{n+1} &= v_y^n - k_2\nabla \times \mathbf{G}\boldsymbol{m}_n && (5.99) \end{aligned}$$

The corresponding L^2 functional is:

$$\begin{aligned} \mathcal{J}(\boldsymbol{m}, v_x, v_y) = \frac{1}{2}\big(&\|\boldsymbol{m}_{n+1} - k_1\mathbf{G}\nabla v_x^{n+1} - k_1\mathbf{G}\nabla \times v_y^{n+1} - \boldsymbol{f}_1\|_{0,\Omega}^2 \\ &+ \|v_x^{n+1} - k_1\nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1} - f_2\|_{0,\Omega}^2 + \|v_y^{n+1} + k_1\nabla \times \mathbf{G}\boldsymbol{m}_{n+1} - f_3\|_{0,\Omega}^2\big) \end{aligned} \tag{5.100}$$

with $\boldsymbol{f}_1 = \partial \boldsymbol{m}_n + k_2 \mathbf{G}\nabla v_x + k_2 \mathbf{G}\nabla \times v_y$, $f_2 = v_x^n + k_2 \nabla \cdot \mathbf{G}\boldsymbol{m}_n$, $f_3 = v_y^n - k_2 \nabla \times \mathbf{G}\boldsymbol{m}_n$.

Now the following theorem establishes an energy estimate, which implies stability of the numerical method corresponding to functional Eq. (5.100).

Theorem 7 *Assuming that the a solution $\boldsymbol{m}_N, v_x^N, v_y^N$ was found by using N times the numerical method based on the the bilinear form stemming from functional Eq. (5.100). Then the following estimate holds for $\theta = 0.5$ with initial conditions $(v_x^0, v_y^0, \boldsymbol{m}) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)^2$:*

$$\begin{aligned}
&\|v_x^{N+1}\|_{0,\Omega}^2 + k_1^2 \|\mathbf{G}\nabla v_x^{N+1}\|_{0,\Omega}^2 + \|v_y^{N+1}\|_{0,\Omega}^2 + k_1^2 \|\mathbf{G}\nabla \times v_y^{N+1}\|_{0,\Omega}^2 + \\
&\|\boldsymbol{m}_{N+1}\|_{0,\Omega}^2 + k_1^2 |\mathbf{G}\boldsymbol{m}_{N+1}|_{1,\Omega}^2 + 2k_1^2 (\mathbf{G}\nabla v_x^{N+1}, \mathbf{G}\nabla \times v_y^{N+1})_{0,\Omega} \leq \\
&\qquad \|v_x^0\|_{0,\Omega}^2 + k_1^2 \|\mathbf{G}\nabla v_x^0\|_{0,\Omega}^2 + \|v_y^0\|_{0,\Omega}^2 + k_1^2 \|\mathbf{G}\nabla \times v_y^0\|_{0,\Omega}^2 + \\
&\qquad \|\boldsymbol{m}_0\|_{0,\Omega}^2 + k_1^2 |\mathbf{G}\boldsymbol{m}_0|_{1,\Omega}^2 + +2k_1^2 (\mathbf{G}\nabla v_x^0, \mathbf{G}\nabla \times v_y^0)_{0,\Omega} + \\
&2k_1 \sum_{n=0}^{N} (\langle \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m}_{n+1}, v_x^{n+1} \rangle_\Gamma + \langle \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m}_n, v_x^n \rangle_\Gamma - \\
&\qquad \langle \boldsymbol{n} \times \mathbf{G}\boldsymbol{m}_n, v_y^n \rangle_\Gamma - \langle \boldsymbol{n} \times \mathbf{G}\boldsymbol{m}_{n+1}, v_y^{n+1} \rangle_\Gamma)
\end{aligned} \tag{5.101}$$

Proof: Applying lemma 3 to the bilinear form which corresponds to functional Eq. (5.100) leads to the following estimate:

$$\begin{aligned}
&\|\boldsymbol{m}_{n+1} - k_1 \mathbf{G}\nabla v_x^{n+1} - k_1 \mathbf{G}\nabla \times v_y^{n+1}\|_{0,\Omega}^2 + \|v_x^{n+1} - k_1 \nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + \\
&\qquad \|v_y^{n+1} - k_1 \nabla \times \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 \leq \\
&\|\boldsymbol{m}_n + k_2 \mathbf{G}\nabla v_x^n + k_2 \mathbf{G}\nabla \times v_y^n\|_{0,\Omega}^2 + \|v_x^n + k_2 \nabla \cdot \mathbf{G}\boldsymbol{m}_n\|_{0,\Omega}^2 + \\
&\qquad \|v_y^n + k_2 \nabla \times \mathbf{G}\boldsymbol{m}_n\|_{0,\Omega}^2
\end{aligned} \tag{5.102}$$

Again the bound for the left hand side will be shown first. Multiplying the terms of the left hand side out leads to:

$$\begin{aligned}
&\|\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + \|k_1 \mathbf{G}\nabla v_x^{n+1}\|_{0,\Omega}^2 + \|k_1 \mathbf{G}\nabla \times v_y^{n+1}\|_{0,\Omega}^2 \\
&- 2k_1 (\boldsymbol{m}_{n+1}, \mathbf{G}\nabla v_x^{n+1})_{0,\Omega} - 2k_1 (\boldsymbol{m}_{n+1}, \mathbf{G}\nabla \times v_y^{n+1})_{0,\Omega} + 2k_1^2 (\mathbf{G}\nabla v_x^{n+1}, \mathbf{G}\nabla \times v_y^{n+1})_{0,\Omega} \\
&+ \|v_x^{n+1}\|_{0,\Omega}^2 + \|k_1 \nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 - 2k_1 (v_x^{n+1}, \nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1})_{0,\Omega} \\
&+ \|v_y^{n+1}\|_{0,\Omega}^2 + \|k_1 \nabla \times \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + 2k_1 (v_y^{n+1}, \nabla \times \mathbf{G}\boldsymbol{m}_{n+1})_{0,\Omega}
\end{aligned} \tag{5.103}$$

By using Green's formulas, most of the domain integrals can be substituted by boundary integrals:

$$\begin{aligned}
&\|\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + \|k_1 \mathbf{G}\nabla v_x^{n+1}\|_{0,\Omega}^2 + \|k_1 \mathbf{G}\nabla \times v_y^{n+1}\|_{0,\Omega}^2 + 2k_1^2 (\mathbf{G}\nabla v_x^{n+1}, \mathbf{G}\nabla \times v_y^{n+1})_{0,\Omega} \\
&+ \|v_x^{n+1}\|_{0,\Omega}^2 + \|k_1 \nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 - 2k_1 \langle \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m}_{n+1}, v_x^{n+1} \rangle_\Gamma \\
&+ \|v_y^{n+1}\|_{0,\Omega}^2 + \|k_1 \nabla \times \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + 2k_1 \langle \boldsymbol{n} \times \mathbf{G}\boldsymbol{m}_{n+1}, v_y^{n+1} \rangle_\Gamma
\end{aligned} \tag{5.104}$$

Sorting the different terms by their primary variable, the left hand side term becomes:

$$\begin{aligned}
&\|v_x^{n+1}\|_{0,\Omega}^2 + k_1^2\|\mathbf{G}\nabla v_x^{n+1}\|_{0,\Omega}^2 + \|v_y^{n+1}\|_{0,\Omega}^2 + k_1^2\|\mathbf{G}\nabla \times v_y^{n+1}\|_{0,\Omega}^2 + \\
&\|\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + k_1^2\|\nabla \cdot \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 + k_1^2\|\nabla \times \mathbf{G}\boldsymbol{m}_{n+1}\|_{0,\Omega}^2 \\
&- 2k_1\langle \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m}_{n+1}, v_x^{n+1}\rangle_\Gamma + 2k_1\langle \boldsymbol{n} \times \mathbf{G}\boldsymbol{m}_{n+1}, v_y^{n+1}\rangle_\Gamma \\
&+ 2k_1^2(\mathbf{G}\nabla v_x^{n+1}, \mathbf{G}\nabla \times v_y^{n+1})_{0,\Omega}
\end{aligned} \tag{5.105}$$

For the scalars v_x and v_y the H^1 equivalence is straightforward while it follows for the vector unknown $\boldsymbol{m}$ from Lemma 4, if the boundary conditions Eqs. (5.93–5.96) are assumed to hold.

For the right hand side, the same steps lead to the following expression:

$$\begin{aligned}
&\|v_x^{n}\|_{0,\Omega}^2 + k_2^2\|\mathbf{G}\nabla v_x^{n}\|_{0,\Omega}^2 + \|v_y^{n}\|_{0,\Omega}^2 + k_2^2\|\mathbf{G}\nabla \times v_y^{n}\|_{0,\Omega}^2 + \\
&\|\boldsymbol{m}_{n}\|_{0,\Omega}^2 + k_2^2\|\nabla \cdot \mathbf{G}\boldsymbol{m}_{n}\|_{0,\Omega}^2 + k_2^2\|\nabla \times \mathbf{G}\boldsymbol{m}_{n}\|_{0,\Omega}^2 \\
&+ 2k_2\langle \boldsymbol{n} \cdot \mathbf{G}\boldsymbol{m}_{n}, v_x^{n}\rangle_\Gamma - 2k_2\langle \boldsymbol{n} \times \mathbf{G}\boldsymbol{m}_{n}, v_y^{n}\rangle_\Gamma + 2k_2^2(\mathbf{G}\nabla v_x^{n}, \mathbf{G}\nabla \times v_y^{n})_{0,\Omega}
\end{aligned} \tag{5.106}$$

Summing these expressions for the left- and right hand side from 1 to N shows the theorem. □

Interestingly, the stability of the vm-formulation does not require velocity boundary conditions on the some parts of the boundary. But on the other hand pure velocity boundary conditions are not sufficient for stability as the conditions of Lemma 4 must always be satisfied to obtain H^1 stability.

Remark: Although stability of the augmented least squares formulation for the 2D wave equation was shown in the previous section, it was not directly clear that the corresponding system for one timestep is elliptic. In 2D the system had 4 equations but only 3 unknowns. This problem can be overcome by adding a slack variable ϑ, which is constant and does not go into computations (cf. [83]):

$$\frac{\partial \boldsymbol{p}}{\partial t} - c\nabla v - c\nabla \times \vartheta = 0 \quad \text{in} \quad \Omega \tag{5.107}$$

$$\frac{\partial v}{\partial t} - \nabla \cdot c\boldsymbol{p} = 0 \quad \text{in} \quad \Omega \tag{5.108}$$

$$\nabla \times c\boldsymbol{p} = 0 \quad \text{in} \quad \Omega. \tag{5.109}$$

The principal part of this system is up to some multiplicative constants exactly the same as for the vm-formulation .

5.6 Stability of More General Least Squares Systems

The first order systems which were shown to be stable in the last section, had a symmetric structure. This property allowed the elimination of the derivative terms and thus was

the cornerstone of showing stability. But the next theorem will establish stability under less restrictive conditions.

Theorem 8 *Let $\boldsymbol{u}_0 \in \mathbf{V}(\Omega)$ where $\mathbf{V}(\Omega)$ is a Hilbert space defined over a domain Ω. Then assuming continuity and coercivity of the bilinear form which corresponds to the left hand side operator and assuming continuity of the bilinear form corresponding to the right hand side operator:*

$$\frac{1}{c_1}\|\boldsymbol{u}\|^2_{\mathbf{V},\Omega} \leq \quad (\mathcal{E}_l\boldsymbol{u}, \mathcal{E}_l\boldsymbol{u})_{0,\Omega} \quad \leq c_1\|\boldsymbol{u}\|^2_{\mathbf{V},\Omega} \quad \forall \boldsymbol{u} \in \mathbf{V}(\Omega) \tag{5.110}$$

$$(\mathcal{E}_r\boldsymbol{u}, \mathcal{E}_r\boldsymbol{v})_{0,\Omega} \quad \leq c_2\|\boldsymbol{u}\|_{\mathbf{V},\Omega} \cdot \|\boldsymbol{v}\|_{\mathbf{V},\Omega} \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbf{V}(\Omega). \tag{5.111}$$

there exists a positive constant c_3 independent of N, such that the solution obtained by iterating N timesteps satisfies:

$$\frac{1}{c_3}\|\boldsymbol{u}_{N+1}\|_{\mathbf{V},\Omega} \leq c_3\|\boldsymbol{u}_0\|_{\mathbf{V},\Omega} + \Delta t \sum_{n=1}^{N} \left\| \frac{\partial \boldsymbol{u}_n}{\partial t} \right\|_{0,\Omega} \tag{5.112}$$

Proof: The continuity and coercivity of the bilinear form corresponding to the left hand side operator $\mathcal{E}_l$ implies the existence and uniqueness of a solution $\boldsymbol{u}_{n+1}$ for each $\boldsymbol{u}_n$. By definition this solution $\boldsymbol{u}_{n+1}$ satisfies the following equation:

$$(\mathcal{I}\boldsymbol{u}_{n+1} + k_1\mathcal{L}\boldsymbol{u}_{n+1}, \mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v})_{0,\Omega} = (\mathcal{I}\boldsymbol{u}_n - k_2\mathcal{L}\boldsymbol{u}_n, \mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v})_{0,\Omega} \quad \forall \boldsymbol{v} \in \mathbf{V}(\Omega) \tag{5.113}$$

Thus also:

$$\sum_{n=0}^{N} (\mathcal{I}\boldsymbol{u}_{n+1} + k_1\mathcal{L}\boldsymbol{u}_{n+1} - \mathcal{I}\boldsymbol{u}_n + k_2\mathcal{L}\boldsymbol{u}_n, \mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v})_{0,\Omega} = 0 \quad \forall \boldsymbol{v} \in \mathbf{V}(\Omega) \tag{5.114}$$

This expression can be simplified to

$$\begin{aligned}(\mathcal{I}\boldsymbol{u}_{N+1} + k_1\mathcal{L}\boldsymbol{u}_{N+1}, \mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v})_0 &= (\mathcal{I}\boldsymbol{u}_0 - k_2\mathcal{L}\boldsymbol{u}_0, \mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v})_0 \\ &\quad + \Delta t \sum_{n=1}^{N} (-\mathcal{L}\boldsymbol{u}_n, \mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v})_0 \quad \forall \boldsymbol{v} \in \mathbf{V}(\Omega)\end{aligned} \tag{5.115}$$

Setting $\boldsymbol{v} = \boldsymbol{u}_{N+1}$, using the Cauchy-Schwarz inequality and dividing by $\|\mathcal{I}\boldsymbol{v} + k_1\mathcal{L}\boldsymbol{v}\|_{0,\Omega}$ gives a first result:

$$\|\mathcal{I}\boldsymbol{u}_{N+1} + k_1\mathcal{L}\boldsymbol{u}_{N+1}\|_{0,\Omega} \leq \|\mathcal{I}\boldsymbol{u}_0 - k_2\mathcal{L}\boldsymbol{u}_0\|_{0,\Omega} + \Delta t \sum_{n=1}^{N} \|\mathcal{L}\boldsymbol{u}_n\|_{0,\Omega} \tag{5.116}$$

As the partial differential equation was

$$\frac{\partial \boldsymbol{u}}{\partial t} + \mathcal{L}\boldsymbol{u} = 0, \tag{5.117}$$

clearly

$$\frac{\partial \boldsymbol{u}}{\partial t} = -\mathcal{L}\boldsymbol{u}. \tag{5.118}$$

This holds also at the discrete time instances $t = \Delta t n$. Thus

$$\frac{\partial \boldsymbol{u}_n}{\partial t} = -\mathcal{L}\boldsymbol{u}_n, \tag{5.119}$$

and hence

$$\|\mathcal{I}\boldsymbol{u}_{N+1} + k_1\mathcal{L}\boldsymbol{u}_{N+1}\|_{0,\Omega} \leq \|\mathcal{I}\boldsymbol{u}_0 - k_2\mathcal{L}\boldsymbol{u}_0\|_{0,\Omega} + \Delta t \sum_{n=1}^{N} \left\| \frac{\partial \boldsymbol{u}_n}{\partial t} \right\|_{0,\Omega}. \tag{5.120}$$

Together with Eqs. (5.110–5.111) this shows the theorem. □

Remark: With this theorem stability could probably be shown under less restrictive conditions for some of the previously introduced first order formulations. But these estimates would then be less sharp and the main objective of this chapter is to find a working first order formulation for linear elastodynamics. Therefore the application of this theorem to the previously mentioned problems is not illustrated anymore as it would not lead to fundamentally new results.

5.7 The *vH*-Formulation for Transient Elasticity Problems

Section 5.2 gave a short review of least squares methods for the solution of the equations of linear elasticity in the stationary case. Inspired by the promising results which were obtained with the formulation based on the displacement gradient (cf. [44], [88] and more), this was selected as a basis for further research. As new unknown, the temporal derivative of the displacement vector was chosen. Thus the resulting formulation has six unknowns in 2D. These represent the spatial and temporal derivatives of the displacement vector $\boldsymbol{u} = (u_x, u_y)^T$. To get a more compact writing, a notation similar to the one introduced in [44] will be used. It introduces the displacement gradient $\mathbf{H} = \nabla \boldsymbol{u}$ (in vector notation):

$$\mathbf{H} = (H_{11}, H_{12}, H_{21}, H_{22})^T = \left(\frac{\partial u_x}{\partial x}, \frac{\partial u_x}{\partial y}, \frac{\partial u_y}{\partial x}, \frac{\partial u_y}{\partial y} \right)^T \tag{5.121}$$

and the following fourth order tensor (written as matrix):

$$\mathbb{A} = \begin{pmatrix} c_1 & 0 & 0 & c_3 \\ 0 & c_2 & c_2 & 0 \\ 0 & c_2 & c_2 & 0 \\ c_3 & 0 & 0 & c_1 \end{pmatrix} \tag{5.122}$$

with $c_1 = (\lambda + 2\mu_s)/\rho_s$, $c_2 = \mu_/\rho_s$ and $c_3 = \lambda/\rho_s$. Furthermore the velocity vector $\boldsymbol{v}$ will be required. With this notation a first order system which is equivalent to Eq. (2.44) becomes:

$$\frac{\partial \mathbf{H}}{\partial t} - \nabla \boldsymbol{v} = 0 \quad \text{in} \quad \Omega \tag{5.123}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} - \nabla \cdot \mathbb{A}\mathbf{H} = 0 \quad \text{in} \quad \Omega \tag{5.124}$$

Introducing the usual semidiscretisation in time, based on the class of θ-methods, the following evolution equation can be obtained:

$$\mathbf{H}_{n+1} - k_1 \nabla \boldsymbol{v}_{n+1} = \mathbf{H}_n + k_2 \nabla \boldsymbol{v}_n \quad \text{in} \quad \Omega \tag{5.125}$$

$$\boldsymbol{v}_{n+1} - k_1 \nabla \cdot \mathbb{A}\mathbf{H}_{n+1} = \boldsymbol{v}_n + k_2 \nabla \cdot \mathbb{A}\mathbf{H}_n \quad \text{in} \quad \Omega \tag{5.126}$$

Clearly the boundary conditions have to be consistent from a physical point of view (cf. p.102 in [74]). Hence either the *geometric boundary conditions* or the *dynamic boundary conditions* have to be specified on each point of Γ (cf. [74]). The geometric boundary conditions are either the displacements or the velocities, while the dynamic boundary conditions refer to the surface forces. Combinations of these are admissible if the boundary conditions do not contradict. An example is the gliding of a material body along a rigid wall, where the normal component of the velocity and the tangential stress can be prescribed (cf. [74]). In the following sections, the following boundary conditions will be assumed to be necessary and sufficient for a well posed problem:

$$\boldsymbol{v} = 0 \quad \text{on} \quad \Gamma_D \tag{5.127}$$

$$\boldsymbol{n} \cdot \mathbb{A}\mathbf{H} = 0 \quad \text{on} \quad \Gamma_N \tag{5.128}$$

where Γ_D and Γ_N denote two non overlapping parts of the boundary Γ of the domain Ω. For simplicity it will be assumed that $\Gamma_D \neq \emptyset$.

5.7.1 The Unstabilised vH-Formulation

First the direct application of the LSFEM idea to Eqs. (5.125–5.126 will be considered. It leads to the following functional, which will be minimised in each time step:

$$\begin{aligned}\mathcal{J}(\mathbf{H}_{n+1}, \boldsymbol{v}_{n+1}, \boldsymbol{f}) = \frac{1}{2}(&\|\mathbf{H}_{n+1} - k_1 \nabla \boldsymbol{v}_{n+1} - \boldsymbol{f}_1\|^2_{0,\Omega} + \\ &\|\boldsymbol{v}_{n+1} - k_1 \nabla \cdot \mathbb{A}\mathbf{H}_{n+1} - \boldsymbol{f}_2\|^2_{0,\Omega})\end{aligned} \tag{5.129}$$

with $\boldsymbol{f}_1 = \mathbf{H}_n + k_2 \nabla \boldsymbol{v}_n$, $\boldsymbol{f}_2 = \boldsymbol{v}_n + k_2 \nabla \cdot \mathbb{A}\mathbf{H}_n$ and $\boldsymbol{f} = (\boldsymbol{f}_1^T, \boldsymbol{f}_2^T)^T$.

Unfortunately, this formulation is unconditionally unstable in conjunction with the standard finite element basis and boundary conditions Eqs. (5.127–5.128) as the next theorem shows:

Theorem 9 *The least squares finite element method based on functional Eq. (5.129) is not coercive in the H^1 norm. Hence it is not suited for solving Eq. (2.44) with boundary conditions Eqs. (5.127- 5.128).*

Proof: Using a simple counterexample it will be demonstrated that the bilinear form corresponding to Eq. (5.129) is not H^1-coercive with boundary conditions Eqs. (5.127-5.128). This implies that the problem which is solved in each timestep is not well posed in the selected spaces, and hence the method is inherently unstable when used with finite dimensional subspaces of $H^1(\Omega)$.

The counterexample is similar to the one presented on p.809 in [19] for the pure L^2 $up\omega$-formulation of the Stokes equations. A unit square will serve as domain Ω. Setting $H_{12} = -\cos(nx)\exp(ny)$, $H_{21} = \cos(nx)\exp(ny)$, $v_x = v_y = H_{11} = H_{22} = 0$ the value of the functional Eq. (5.129) is (computed with MuPAD):

$$\frac{1}{n^2}\left(\frac{n(e^n)^2}{4} + \frac{(e^n)^2\sin(2n)}{8} - \left(\frac{n}{4} + \frac{\sin(2n)}{8}\right)\right) \approx O\left(\frac{1}{n}(e^n)^2\right). \tag{5.130}$$

For the H^1 norm of H_{12} the following result can be found (with MuPAD):

$$\begin{aligned}\|H_{12}\|^2_{1,\Omega} =& \frac{1}{n^2}\left(\frac{n(e^n)^2}{4} - \frac{n^2(e^n)^2}{4} + \frac{n^3(e^n)^2}{2} + \frac{(e^n)^2\sin(2n)}{8} + \frac{n^2(e^n)^2\cos(2n)}{4}\right) - \\ & \frac{1}{n^2}\left(\frac{n}{4} - \frac{n^2}{4} + \frac{n^3}{2} + \frac{\sin(2n)}{8} + \frac{n^2\cos(2n)}{4}\right) \approx O\left(n(e^n)^2\right)\end{aligned} \tag{5.131}$$

As $H_{21} = -H_{12}$, $\|H_{21}\|_1^2 = \|H_{12}\|_1^2$. Further $\|H_{11}\|_1^2 = \|H_{22}\|_1^2 = \|v_x\|_1^2 = \|v_y\|_1^2 = 0$. Hence

$$C(\|\mathbf{H}\|_1^2 + \|\boldsymbol{v}\|_1^2) \approx O\left(n(e^n)^2\right) \geq \mathcal{J}(\mathbf{H}, \boldsymbol{v}, 0) \approx O\left(\frac{1}{n}(e^n)^2\right). \tag{5.132}$$

which is a contradiction to H^1-coercivity. $\square$.

Remark: Taking a look onto the principal part of Eqs. (5.125–5.126)

$$\mathcal{L}^p(\xi,\eta) = \begin{pmatrix} 0 & 0 & 0 & 0 & \xi & 0 \\ 0 & 0 & 0 & 0 & \eta & 0 \\ 0 & 0 & 0 & 0 & 0 & \xi \\ 0 & 0 & 0 & 0 & 0 & \eta \\ c_1\xi & c_2\eta & c_2\eta & c_3\xi & 0 & 0 \\ c_3\eta & c_2\xi & c_2\xi & c_1\eta & 0 & 0 \end{pmatrix} \tag{5.133}$$

with determinant:

$$|\mathcal{L}^p| = 0 \tag{5.134}$$

shows that the system of equations is not elliptic. Although that does not automatically indicate that a least squares formulation does not work properly as the example of the 2D wave equation showed, it gives a first hint that the formulation might be troublesome. Furthermore the unstabilised least squares formulation for the 2D wave equation was only well posed in the $H(\mathrm{div})$ norm.

Augmenting the 2D wave equation by a seemingly redundant equation restricts the solution space to $H^1 \times (H^1)^2$. Hence this technique could also work for the equations of linear elasticity.

5.7.2 The Augmented vH-Formulation

Similar to the stabilisation term for wave equation, the Eqs. (5.125–5.126) will be augmented by the equation $\nabla \times \mathbf{H}$ which leads to the following system of equations:

$$\mathbf{H}_{n+1} - k_1 \nabla \boldsymbol{v}_{n+1} = \mathbf{H}_n + k_2 \nabla \boldsymbol{v}_n \quad \text{in} \quad \Omega \tag{5.135}$$

$$\boldsymbol{v}_{n+1} - k_1 \nabla \cdot \mathbb{A}\mathbf{H}_{n+1} = \boldsymbol{v}_n + k_2 \nabla \cdot \mathbb{A}\mathbf{H}_n \quad \text{in} \quad \Omega \tag{5.136}$$

$$k_1 \nabla \times \mathbf{H}_{n+1} = k_2 \nabla \times \mathbf{H}_n \quad \text{in} \quad \Omega \tag{5.137}$$

As this system has 8 equations but only 6 unknowns, it cannot be elliptic. Introducing two additional slack variables ϑ_1 and ϑ_2 which form the vector $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)^T$ leads to a properly defined elliptic system. These slack variables are only required for the analysis and will not appear in computations. This approach is similar to the one proposed in [83] for first order formulation of the 2D diffusion equation or in [21] for the $up\omega$-formulation of the Stokes equations in 3D. The rotation of the slack variables is added to Eq. (5.135):

$$\mathbf{H}_{n+1} - k_1(\nabla \boldsymbol{v}_{n+1} + \nabla \times \boldsymbol{\vartheta}_{n+1}) = \mathbf{H}_n + k_2(\nabla \boldsymbol{v}_n + \nabla \times \boldsymbol{\vartheta}_n) \quad \text{in} \quad \Omega \tag{5.138}$$

Here $\nabla\times$ is meant componentwise, $\nabla \times \boldsymbol{\vartheta} = (\frac{\vartheta_1}{\partial y}, -\frac{\vartheta_1}{\partial x}, \frac{\vartheta_2}{\partial y}, -\frac{\vartheta_2}{\partial x})^T$. With this modification the system has eight equations and eight unknowns and is elliptic of order 4. A detailed proof of ellipticity will follow in subsection 5.7.3. Looking at Eq. (5.150), which shows the principal part of the system, the block diagonal structure is obvious. This structure reveals that the system does consist of three decoupled elliptic systems. The first block is of order two and hence must be accompanied by two boundary conditions, while the other two blocks are of order one and must be supported by one boundary condition. Therefore it is not possible to impose four boundary conditions on the large 4×4 block and none on the other two blocks.

The corresponding least squares functional is:

$$\begin{aligned} \mathcal{J}(\mathbf{H}_{n+1}, \boldsymbol{v}_{n+1}, \boldsymbol{f}) = \frac{1}{2} \big(& \|\mathbf{H}_{n+1} - k_1 \nabla \boldsymbol{v}_{n+1} - \boldsymbol{f}_1\|^2_{0,\Omega} + \\ & \|\boldsymbol{v}_{n+1} - k_1 \nabla \cdot \mathbb{A}\mathbf{H}_{n+1} - \boldsymbol{f}_2\|^2_{0,\Omega} + \\ & \|k_1 \nabla \times \mathbf{H}_{n+1} - \boldsymbol{f}_3\|^2_{0,\Omega} \big) \end{aligned} \tag{5.139}$$

with the following right hand side terms $\boldsymbol{f}_1 = \mathbf{H}_n + k_2 \nabla \boldsymbol{v}_n$, $\boldsymbol{f}_2 = \boldsymbol{v}_n + k_2 \nabla \cdot \mathbb{A}\mathbf{H}_n$, $\boldsymbol{f}_3 = k_2 \nabla \times \mathbf{H}_n$ and $\boldsymbol{f} = (\boldsymbol{f}_1^T, \boldsymbol{f}_2^T, \boldsymbol{f}_3^T)^T$.

5.7.3 Coercivity

In this section the a priori estimates for the functional corresponding to the augmented formulation will be derived using elliptic theory and a modification of the proof in [135]. Adapting the results for the stationary formulation from literature (cf. [44] and [88]) seems to be possible as well but is more difficult.

For the analysis two different boundary conditions will be considered. These correspond to Eqs. (5.127–5.128), which are common in structural mechanics. The Neumann boundary condition, which defines the normal tractions on the boundary takes the following form in the vH-formulation :

$$\begin{aligned} n_x(c_1 H_{11} + c_3 H_{22}) + n_y(c_2 H_{12} + c_2 H_{21}) &= 0 \quad \text{on} \quad \Gamma_N && (5.140)\\ n_x(c_2 H_{12} + c_2 H_{21}) + n_y(c_3 H_{11} + c_1 H_{22}) &= 0 \quad \text{on} \quad \Gamma_N && (5.141)\\ \vartheta_1 &= 0 \quad \text{on} \quad \Gamma_N && (5.142)\\ \vartheta_2 &= 0 \quad \text{on} \quad \Gamma_N && (5.143) \end{aligned}$$

Equivalently, the Dirichlet boundary conditions, which corresponds to a fixed end become:

$$\begin{aligned} n_y H_{11} - n_x H_{12} &= 0 \quad \text{on} \quad \Gamma_D && (5.144)\\ n_y H_{21} - n_x H_{22} &= 0 \quad \text{on} \quad \Gamma_D && (5.145)\\ v_x &= 0 \quad \text{on} \quad \Gamma_D && (5.146)\\ v_y &= 0 \quad \text{on} \quad \Gamma_D && (5.147) \end{aligned}$$

These boundary conditions satisfy the previously required conditions that each block of the principal part is accompanied by a sufficiently large number of boundary conditions.

Theorem 10 *Let $\mathcal{J}(\mathbf{H}, \boldsymbol{v}, \boldsymbol{f})$ be the functional defined in Eq. (5.139). Then there exists a positive constant c_1 such that*

$$\frac{1}{c_1}(\|\mathbf{H}\|_{1,\Omega}^2 + \|\boldsymbol{v}\|_{1,\Omega}^2) \leq \mathcal{J}(\mathbf{H}, \boldsymbol{v}, 0) \leq c_1(\|\mathbf{H}\|_{1,\Omega}^2 + \|\boldsymbol{v}\|_{1,\Omega}^2) \tag{5.148}$$

for all $(\mathbf{H}, \boldsymbol{v}) \in H^1(\Omega)^4 \times H^1(\Omega)^2$ which satisfy the boundary conditions Eqs. (5.140–5.147).

Proof: The proof follows the outline of [135]. First it will be shown that the first order system is elliptic in the sense of Petrovski.

The ordering of the unknowns will be as follows:

$$\boldsymbol{w} = (H_{11}, H_{12}, H_{21}, H_{22}, v_x, \vartheta_1, v_y, \vartheta_2)^T \tag{5.149}$$

The principal part of this augmented system is:

$$\mathcal{L}^p(\xi,\eta) = k_1 \left(\begin{array}{cccc|cccc} c_1\xi & c_2\eta & c_2\eta & c_3\xi & & & & \\ c_3\eta & c_2\xi & c_2\xi & c_1\xi & & & & \\ -\eta & \xi & & & & & & \\ & & -\eta & \xi & & & & \\ \hline & & & & \xi & -\eta & & \\ & & & & \eta & \xi & & \\ & & & & & & \xi & -\eta \\ & & & & & & \eta & \xi \end{array}\right) \tag{5.150}$$

with determinant:

$$|\mathcal{L}^p| = -k_1\mu_s(\lambda + 2\mu_s)(\xi^2 + \eta^2)^4. \tag{5.151}$$

The determinant is nonzero for all nonzero real pairs (ξ, η). Thus the system is uniformly elliptic in the sense of Petrovski with order 4.

The boundary conditions will be represented by the following two matrices:

$$\mathbf{C}_1 = \left(\begin{array}{cccccccc} n_x\,(2\,\mu_s + \lambda) & \mu_s\,n_y & \mu_s\,n_y & n_x\,\lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ n_y\,\lambda & \mu_s\,n_x & \mu_s\,n_x & n_y\,(2\,\mu_s + \lambda) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \tag{5.152}$$

for the traction boundary condition and

$$\mathbf{C}_2 = \left(\begin{array}{cccccccc} n_y & -n_x & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & n_y & -n_x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right) \tag{5.153}$$

for the Dirichlet boundary condition.

Setting $(\xi, \eta) = (1,0)$ the coefficients belonging to the two derivative directions are:

$$\mathbf{A} = \left(\begin{array}{cccc|cccc} 2\,\mu_s + \lambda & 0 & 0 & \lambda & & & & \\ 0 & \mu_s & \mu_s & 0 & & & & \\ 0 & 1 & 0 & 0 & & & & \\ 0 & 0 & 0 & 1 & & & & \\ \hline & & & & 1 & 0 & & \\ & & & & 0 & 1 & & \\ & & & & & & 1 & 0 \\ & & & & & & 0 & 1 \end{array}\right) \tag{5.154}$$

and

$$\mathbf{B} = \left(\begin{array}{cccc|cccc} 0 & \mu_s & \mu_s & 0 & & & & \\ \lambda & 0 & 0 & 2\,\mu_s+\lambda & & & & \\ -1 & 0 & 0 & 0 & & & & \\ 0 & 0 & -1 & 0 & & & & \\ \hline & & & & 0 & -1 & & \\ & & & & 1 & 0 & & \\ & & & & & & 0 & -1 \\ & & & & & & 1 & 0 \end{array}\right). \tag{5.155}$$

Using the fact that $\mathbf{A}$ is nonsingular, the inverse $\mathbf{A}^{-1}$ can be found:

$$\mathbf{A}^{-1} = \left(\begin{array}{cccc|cccc} \frac{1}{2\,\mu_s+\lambda} & 0 & 0 & -\frac{\lambda}{2\,\mu_s+\lambda} & & & & \\ 0 & 0 & 1 & 0 & & & & \\ 0 & \frac{1}{\mu_s} & -1 & 0 & & & & \\ 0 & 0 & 0 & 1 & & & & \\ \hline & & & & 1 & 0 & & \\ & & & & 0 & 1 & & \\ & & & & & & 1 & 0 \\ & & & & & & 0 & 1 \end{array}\right) \tag{5.156}$$

With this matrix, the system of differential equations can be transformed into:

$$\frac{\partial \boldsymbol{w}}{\partial x} + \mathbf{A}^{-1}\mathbf{B}\frac{\partial \boldsymbol{w}}{\partial y} = 0 \tag{5.157}$$

with

$$\tilde{\mathbf{B}} = \mathbf{A}^{-1}\mathbf{B} = \left(\begin{array}{cccc|cccc} 0 & \frac{\mu_s}{2\,\mu_s+\lambda} & \frac{\mu_s}{(2\,\mu_s+\lambda)} + \frac{\lambda}{(2\,\mu_s+\lambda)} & 0 & & & & \\ -1 & 0 & 0 & 0 & & & & \\ \frac{\lambda}{\mu_s}+1 & 0 & 0 & \frac{2\,\mu_s+\lambda}{\mu_s} & & & & \\ 0 & 0 & -1 & 0 & & & & \\ \hline & & & & 0 & -1 & & \\ & & & & 1 & 0 & & \\ & & & & & & 0 & -1 \\ & & & & & & 1 & 0 \end{array}\right) \tag{5.158}$$

This matrix has two eigenvalues, i and $-i$ with multiplicities of four. For the eigenvalue $\tau_+ = i$ the following chain of generalised eigenvectors can be found:

$$\boldsymbol{p}_1 = \left(\frac{i\,\mu_s + i\,\lambda}{4\,\mu_s}, \frac{\mu_s + \lambda}{4\,\mu_s}, \frac{\mu_s + \lambda}{4\,\mu_s}, \frac{-i\,\mu_s - i\,\lambda}{4\,\mu_s}, 0, 0, 0, 0\right)^T \tag{5.159}$$

$$\boldsymbol{p}_2 = \left(1/2, \frac{-3\,i\,\mu_s - i\,\lambda}{4\,\mu_s}, \frac{i\,\mu_s + i\,\lambda}{4\,\mu_s}, 0, 0, 0, 0, 0\right)^T \tag{5.160}$$

$$\boldsymbol{p}_3 = \left(0,0,0,0,1/2, \frac{i}{2}, 0, 0\right)^T \tag{5.161}$$

$$\boldsymbol{p}_4 = \left(0,0,0,0,0,0,1/2, \frac{i}{2}\right)^T \tag{5.162}$$

The matrix, which can be constructed from these eigenvectors and their conjugates:

$$\mathbf{P} = (\boldsymbol{p}_1, \bar{\boldsymbol{p}}_1, \boldsymbol{p}_2, \bar{\boldsymbol{p}}_2, \boldsymbol{p}_3, \bar{\boldsymbol{p}}_3, \boldsymbol{p}_4, \bar{\boldsymbol{p}}_4)^T \tag{5.163}$$

is nonsingular. The inverse of $\mathbf{P}$ is then

$$\mathbf{L} = \left(\begin{array}{cccc|cccc}
0 & 1 & 0 & 1 & & & & \\
\frac{i\,\mu_s}{2\,i\,\mu_s + i\,\lambda} & \frac{i\,\mu_s}{2\,\mu_s + \lambda} & \frac{i\,\mu_s}{2\,i\,\mu_s + i\,\lambda} & \frac{-i\,\mu_s}{2\,\mu_s + \lambda} & & & & \\
\frac{\mu_s\,\lambda + 3\,\mu_s^2}{3\,\mu_s\,\lambda + 2\,\mu_s^2 + \lambda^2} & -\frac{\mu_s\,(i\,\mu_s + i\,\lambda)}{3\,\mu_s\,\lambda + 2\,\mu_s^2 + \lambda^2} & \frac{\mu_s\,\lambda + 3\,\mu_s^2}{3\,\mu_s\,\lambda + 2\,\mu_s^2 + \lambda^2} & \frac{\mu_s\,(i\,\mu_s + i\,\lambda)}{3\,\mu_s\,\lambda + 2\,\mu_s^2 + \lambda^2} & & & & \\
-\frac{2\,\mu_s}{i\,\mu_s + i\,\lambda} & 1 & -\frac{\mu_s\,(4\,i\,\mu_s + 2\,i\,\lambda)}{3\,\mu_s\,\lambda + 2\,\mu_s^2 + \lambda^2} & 1 & & & & \\
\hline
 & & & & 1 & 1 & & \\
 & & & & -i & i & & \\
 & & & & & & 1 & 1 \\
 & & & & & & -i & i
\end{array}\right) \tag{5.164}$$

Which can be written symbolically as:

$$\mathbf{L} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \bar{\boldsymbol{q}}_1, \bar{\boldsymbol{q}}_2, \boldsymbol{q}_3, \bar{\boldsymbol{q}}_3, \boldsymbol{q}_4, \bar{\boldsymbol{q}}_4) \tag{5.165}$$

Finally, the determinant of the boundary operators multiplied with the matrix $(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4)$ is analysed. For the two boundary conditions

$$\det(\mathbf{C}_1(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4)) = \frac{4\,(-i\,\mu_s\,n_y - n_x\,\lambda)\,(n_x + i\,n_y)^2\,\mu_s^2}{i\,\mu_s\,n_y + n_x\,\lambda} \neq 0 \quad \text{on} \quad \Gamma_N \tag{5.166}$$

and

$$\det(\mathbf{C}_2(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4)) = \frac{(n_y - i\,n_x)^2\,(3\,\mu_s + \lambda)\,\mu_s}{(2\,\mu_s + \lambda)\,(\mu_s + \lambda)} \neq 0 \quad \text{on} \quad \Gamma_D \tag{5.167}$$

hold for $(n_x, n_y) \neq (0,0)$. Hence the Lopatinski condition is fulfilled for the two boundary conditions Eqs. (5.140–5.143) and Eqs. (5.144–5.147). Using lemma 8.2.2 from [129] shows then the a priori estimate. □

5.7.4 Stability

Now the stability of the vH-formulation will be analysed with the same tools which were already used in the previous sections. The basis for the numerical method and the analysis is again the least squares functional Eq. (5.139).

Theorem 11 *Let* $(\mathbf{H}_0, \boldsymbol{v}_0) \in H^1(\Omega)^4 \times H^1(\Omega)^2$. *Then the solution* $(\mathbf{H}_N, \boldsymbol{v}_N)$ *which is found by proceeding* N *time steps with the formulation based on functional Eq. (5.139), satisfies:*

$$\begin{aligned}\frac{1}{c_1}(\|\mathbf{H}_{N+1}\|_{1,\Omega} + \|v_{N+1}\|_{1,\Omega}) \leq & c_1(\|\mathbf{H}_0\|_{1,\Omega} + \|v_0\|_{1,\Omega}) + \\ & \Delta t \sum_{n=1}^{N} \|((\nabla \cdot \mathbf{A}\mathbf{H}_n)^T, (\nabla \boldsymbol{v}_n)^T)^T\|_{0,\Omega}\end{aligned} \tag{5.168}$$

where c_1 *is a positive constant independent of* N.

Proof: The theorem is a consequence of theorem 8 and 10. □

5.8 The $v\sigma$-Formulation for Linear Elastodynamic

Several publications consider the use of the physical stresses as additional unknowns in the first order least squares setting (cf. [37], [40], [38]). The formulation proposed in [37] uses the following set of equations:

$$\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{u}) = 0 \quad \text{in} \quad \Omega \tag{5.169}$$

$$\nabla \cdot \sigma = 0 \quad \text{in} \quad \Omega \tag{5.170}$$

It can easily be transformed into a formulation for the transient problem by formulating the constitutive law in terms of the stress and strain rates:

$$\mathbb{C}^{-1/2}\dot{\sigma} - \mathbb{C}^{1/2}\dot{\varepsilon}(\boldsymbol{u}) = 0 \quad \text{in} \quad \Omega. \tag{5.171}$$

The linearised strain rate can then be expressed by the velocity unknowns:

$$\varepsilon(\boldsymbol{v}) = \frac{1}{2}(\nabla_X \dot{\boldsymbol{u}} + (\nabla_X \dot{\boldsymbol{u}})^T) = \frac{1}{2}(\nabla_X \boldsymbol{v} + (\nabla_X \boldsymbol{v})^T). \tag{5.172}$$

The index X on the nabla operator indicating derivatives in the material configuration will be dropped, as other configurations will not be needed in the following paragraphs. With these expressions, the equivalent set of first order equations for the transient problem Eq. (2.44) reads:

$$\mathbb{C}^{-1/2}\frac{\partial \sigma}{\partial t} - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}) = 0 \quad \text{in} \quad \Omega \tag{5.173}$$

$$\rho_s \frac{\partial \boldsymbol{v}}{\partial t} - \nabla \cdot \sigma = 0 \quad \text{in} \quad \Omega \tag{5.174}$$

This system could be extended by an additional constraint enforcing the symmetry of the stress tensor. But as it will be shown during the analysis this constraint is not necessary. Furthermore boundary conditions will be needed. The boundary of Ω is divided into two parts Γ_N and Γ_D with $\partial\Omega = \Gamma_N \cup \Gamma_D$, $\Gamma_D \cap \Gamma_N = \emptyset$, $\Gamma_D \neq \emptyset$. Under certain conditions $\Gamma_D \neq \emptyset$ is not required. The inclusion of this special situation complicates the analysis and hence is omitted. On these boundaries, the following boundary conditions will be assumed:

$$\boldsymbol{v} = 0 \quad \text{on} \quad \Gamma_D \tag{5.175}$$

$$\boldsymbol{n} \cdot \sigma = 0 \quad \text{on} \quad \Gamma_N \tag{5.176}$$

For the analysis of the time discretised version of these equations, it will be convenient to introduce the elasticty tensor scaled by k_1:

$$\mathbb{C}_t = k_1 \mathbb{C} \tag{5.177}$$

Using the θ-methods for time discretisation, the time discrete version of Eqs. (5.173–5.174) is

$$\mathbb{C}_t^{-1/2}\sigma_{n+1} - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}_{n+1}) = \mathbb{C}_t^{-1/2}\sigma_n + (k_2/\sqrt{k_1})\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}_n) \qquad \text{in} \quad \Omega \tag{5.178}$$

$$(1/\sqrt{k_1})\boldsymbol{v}_{n+1} - \sqrt{k_1}\nabla \cdot \sigma_{n+1} = (1/\sqrt{k_1})\boldsymbol{v}_n + (k_2/\sqrt{k_1})\nabla \cdot \sigma_n \qquad \text{in} \quad \Omega \tag{5.179}$$

The mass density ρ_s was not included, as it does not influence the stability of the formulation. For the computations it is included into Eq. (5.179) (cf. section B.9).

On the basis of this system of equations, several least squares functionals can be derived. First they will include the symmetry constraint $\frac{1}{2}(\sigma - \sigma^T) = 0$ as it will be necessary for the first parts of the analysis. The L^2 functional is:

$$\begin{aligned}\mathcal{J}(\sigma_{n+1}, \boldsymbol{v}_{n+1}, \boldsymbol{f}) = \frac{1}{2}(&\|\mathbb{C}_t^{-1/2}\sigma_{n+1} - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}_{n+1}) - \boldsymbol{f}_1\|_{0,\Omega}^2 + \\ &\|(1/\sqrt{k_1})\boldsymbol{v}_{n+1} - \sqrt{k_1}\nabla \cdot \sigma_{n+1} - \boldsymbol{f}_2\|_{0,\Omega}^2 + \left\|\frac{1}{2}(\sigma - \sigma^T)\right\|_{0,\Omega}^2)\end{aligned} \tag{5.180}$$

with $\boldsymbol{f}_1 = \mathbb{C}_t^{-1/2}\sigma_n + (k_2/\sqrt{k_1})\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}_n)$, $\boldsymbol{f}_2 = (1/\sqrt{k_1})\boldsymbol{v}_n + (k_2/\sqrt{k_1})\nabla \cdot \sigma_n$ and $\boldsymbol{f} = (\boldsymbol{f}_1^T, \boldsymbol{f}_2^T)^T$.

For the analysis the following negative norm functional will be used as well. It could probably be used also for computations, but this path will not be considered any further:

$$\begin{aligned}\mathcal{J}_{-1}(\sigma_{n+1}, \boldsymbol{v}_{n+1}, \boldsymbol{f}) = \frac{1}{2}(&\|\mathbb{C}_t^{-1/2}\sigma_{n+1} - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}_{n+1}) - \boldsymbol{f}_1\|_{0,\Omega}^2 + \\ &\|(1/\sqrt{k_1})\boldsymbol{v}_{n+1} - \sqrt{k_1}\nabla \cdot \sigma_{n+1} - \boldsymbol{f}_2\|_{-1,\Omega}^2 + \left\|\frac{1}{2}(\sigma - \sigma^T)\right\|_{0,\Omega}^2)\end{aligned} \tag{5.181}$$

The definition of $\boldsymbol{f}_1$, $\boldsymbol{f}_2$ and $\boldsymbol{f}$ is equivalent to the one used for $\mathcal{J}$.

5.8.1 Coercivity with Symmetry Constraint

Coercivity of the least squares functionals $\mathcal{J}$ and $\mathcal{J}_{-1}$ in appropriate norms will be shown in two steps. First the boundedness in terms of the following norms M and M_{-1}:

$$M(\boldsymbol{w},\tau) = \|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{w})\|_{0,\Omega}^2 + \|\mathbb{C}_t^{-1/2}\tau\|_{0,\Omega}^2 + \|\nabla\cdot\tau\|_{0,\Omega}^2 \tag{5.182}$$

$$M_{-1}(\boldsymbol{w},\tau) = \|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{w})\|_{0,\Omega}^2 + \|\mathbb{C}_t^{-1/2}\tau\|_{0,\Omega}^2 + \|\nabla\cdot\tau\|_{-1,D,\Omega}^2 \tag{5.183}$$

will be shown. After that the equivalence of these functionals with the norm of $H(\mathrm{div};\Omega)^d\times H^1(\Omega)^d$ is demonstrated.

Theorem 12 *For $k_1 > 0$ there exist positive constants c_1, c_2 independent of λ, such that*

$$\frac{1}{c_1}M(\boldsymbol{w},\tau) \le \mathcal{J}(\tau,\boldsymbol{w},0) \le c_1 M(\boldsymbol{w},\tau) \tag{5.184}$$

and

$$\frac{1}{c_2}M_{-1}(\boldsymbol{w},\tau) \le \mathcal{J}_{-1}(\tau,\boldsymbol{w},0) \le c_2 M_{-1}(\boldsymbol{w},\tau) \tag{5.185}$$

hold for all $(\boldsymbol{w},\tau)\in H_D^1(\Omega)^d\times H_N(\mathrm{div};\Omega)^d$

Proof: The proof is an adaptation of the results presented in [37] to the slightly modified least squares functional for the transient problem. The upper bound can be established by the Cauchy and triangle inequality.

First a few results from [37] will be cited. These will be used in the proof of the theorem. Korn's inequality states (cf. section VI.3 in [23]):

$$\|\boldsymbol{v}\|_{1,\Omega}^2 \le c_1\|\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2 \quad \forall\boldsymbol{v}\in H_D^1(\Omega)^d \tag{5.186}$$

A tensor τ can be decomposed into its symmetric and skew-symmetric part:

$$\tau = \frac{\tau+\tau^T}{2} + \frac{\tau-\tau^T}{2}. \tag{5.187}$$

If $\mathbf{A}$ and $\mathbf{B}$ are symmetric and skew-symmetric tensors, $\mathbf{A}:\mathbf{B}=0$. The linearised strain rate tensor $\varepsilon(\boldsymbol{v})$ is the symmetric part of the displacement rate gradient. Hence

$$\begin{aligned}(\tau,\varepsilon(\boldsymbol{v}))_{0,\Omega} &= \left(\frac{\tau+\tau^T}{2},\varepsilon(\boldsymbol{v})\right)_{0,\Omega} = \left(\frac{\tau+\tau^T}{2},\nabla\boldsymbol{v}\right)_{0,\Omega}\\ &= (\tau,\nabla\boldsymbol{v})_{0,\Omega} - \left(\frac{\tau-\tau^T}{2},\nabla\boldsymbol{v}\right)_{0,\Omega} = -(\nabla\cdot\tau,\boldsymbol{v})_{0,\Omega} - \left(\frac{\tau-\tau^T}{2},\nabla\boldsymbol{v}\right)_{0,\Omega}\end{aligned} \tag{5.188}$$

As usual the difficult part is to bound M_{-1} by the least squares functional $\mathcal{J}_{-1}$. First $\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})$ will be treated

$$\begin{aligned}\|v\|_{0,\Omega}^2+k_1\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2 &= k_1(\mathbb{C}_t\varepsilon(\boldsymbol{v})-\tau,\varepsilon(\boldsymbol{v}))_{0,\Omega}+k_1(\tau,\varepsilon(\boldsymbol{v}))_{0,\Omega}+(\boldsymbol{v},\boldsymbol{v})_{0,\Omega}\\ &= k_1(\mathbb{C}_t^{-1/2}\tau-\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}),\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}))_{0,\Omega}-k_1(\nabla\cdot\tau,\boldsymbol{v})_{0,\Omega}-\\ &\quad k_1\left(\frac{\tau-\tau^T}{2},\nabla v\right)_{0,\Omega}+(\boldsymbol{v},\boldsymbol{v})_{0,\Omega}\\ &= k_1(\mathbb{C}_t^{-1/2}\tau-\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}),\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}))_{0,\Omega}+((1/\sqrt{k_1})\boldsymbol{v}-\sqrt{k_1}\nabla\cdot\tau,\sqrt{k_1}\boldsymbol{v})_{0,\Omega}-\\ &\sqrt{k_1}\left(\frac{\tau-\tau^T}{2},\sqrt{k_1}\nabla\boldsymbol{v}\right)_{0,\Omega}\end{aligned} \tag{5.189}$$

Using the Cauchy-Schwarz inequality leads to

$$\begin{aligned}&\leq k_1\|\mathbb{C}_t^{-1/2}\tau-\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}+\\ &\|(1/\sqrt{k_1})\boldsymbol{v}-\sqrt{k_1}\nabla\cdot\tau\|_{-1,D,\Omega}\|\sqrt{k_1}\boldsymbol{v}\|_{1,\Omega}+\sqrt{k_1}\left\|\frac{\tau-\tau^T}{2}\right\|_{0,\Omega}\|\sqrt{k_1}\nabla\boldsymbol{v}\|_{0,\Omega}\end{aligned} \tag{5.190}$$

Both terms $\|\sqrt{k_1}v\|_{1,\Omega}$ and $\|\sqrt{k_1}\nabla\boldsymbol{v}\|$ can be bounded by $\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|$ using Korn's inequality. Thus:

$$\begin{aligned}&\leq c_2(k_1\|\mathbb{C}_t^{-1/2}\tau-\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}+\\ &\quad\|(1/\sqrt{k_1})\boldsymbol{v}-\sqrt{k_1}\nabla\cdot\tau\|_{-1,D,\Omega}+\sqrt{k_1}\left\|(\tau-\tau^T)/2\right\|_{0,\Omega})\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}\end{aligned} \tag{5.191}$$

As $k_1\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2\leq\|v\|_{0,\Omega}^2+k_1\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2$:

$$\begin{aligned}\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}\leq&c_3(\|\mathbb{C}_t^{-1/2}\tau-\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}+\\ &(1/k_1)\|(1/\sqrt{k_1})\boldsymbol{v}-\sqrt{k_1}\nabla\cdot\tau\|_{-1,D,\Omega}+(1/\sqrt{k_1})\|(\tau-\tau^T)/2\|_{0,\Omega})\end{aligned} \tag{5.192}$$

At this point the factors of the right hand side which depend on the time step size k_1 will be included into the coercivity constants. Hence the resulting coercivity estimates will not be independent of Δt! Squaring both sides leads to a first result:

$$\begin{aligned}\|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2\leq c_4(\|\mathbb{C}_t^{-1/2}\tau-\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}+&\|(1/\sqrt{k_1})\boldsymbol{v}-\sqrt{k_1}\nabla\cdot\tau\|_{-1,D,\Omega}+\\ &\|(\tau-\tau^T)/2\|_{0,\Omega})^2\leq c_5\mathcal{J}_{-1}(\tau,\boldsymbol{v},0)\end{aligned} \tag{5.193}$$

Now the term $\|\mathbb{C}_t^{-1/2}\tau\|_{0,\Omega}^2$ will be bounded from above using the triangle inequality:

$$\|\mathbb{C}_t^{-1/2}\tau\|_{0,\Omega} \leq \|\mathbb{C}_t^{-1/2}\tau - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega} + \|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega} \tag{5.194}$$

and hence

$$\begin{aligned} \|\mathbb{C}_t^{-1/2}\tau\|_{0,\Omega}^2 &\leq (\|\mathbb{C}_t^{-1/2}\tau - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2 + \|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2)^2 \\ &\leq c_6(\|\mathbb{C}_t^{-1/2}\tau - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2 + \|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|_{0,\Omega}^2) \\ &\leq c_7\mathcal{J}_{-1}(\tau, \boldsymbol{v}, 0) \end{aligned} \tag{5.195}$$

For the bound of $\|\nabla \cdot \tau\|_{0,\Omega}$ again the triangle inequality is used:

$$\sqrt{k_1}\|\nabla \cdot \tau\|_{-1,D,\Omega} \leq \|(1/\sqrt{k_1})\boldsymbol{v} - \sqrt{k_1}\nabla \cdot \tau\|_{-1,D,\Omega} + \|(1/\sqrt{k_1})\boldsymbol{v}\|_{-1,D,\Omega} \tag{5.196}$$

and hence

$$\begin{aligned} k_1\|\nabla \cdot \tau\|_{-1,D,\Omega}^2 &\leq (\|(1/\sqrt{k_1})\boldsymbol{v} - \sqrt{k_1}\nabla \cdot \tau\|_{-1,D,\Omega} + \|(1/\sqrt{k_1})\boldsymbol{v}\|_{-1,D,\Omega})^2 \\ &\leq c_8(\|(1/\sqrt{k_1})\boldsymbol{v} - \sqrt{k_1}\nabla \cdot \tau\|_{-1,D,\Omega}^2 + \|(1/\sqrt{k_1})\boldsymbol{v}\|_{-1,D,\Omega}^2) \\ &\leq c_9\mathcal{J}_{-1}(\tau, \boldsymbol{v},0) \end{aligned} \tag{5.197}$$

Equivalently it can be shown that $\|\nabla \cdot \tau\|_{0,\Omega}$ is bounded by $\mathcal{J}(\tau, \boldsymbol{v},0)$. Noting that $\mathcal{J}_{-1}(\tau, \boldsymbol{v},0) \leq c_{10}\mathcal{J}(\tau, \boldsymbol{v},0)$, the proof is complete. □

5.8.2 Coercivity without Symmetry Constraint

As it was shown in [40] the inclusion of the symmetry constraint into the least squares functional is not necessary to establish coercivity and continuity. The following two least squares functionals are the counterparts of $\mathcal{J}$ and $\mathcal{J}_{-1}$ without the symmetry constraint:

$$\begin{aligned} \mathcal{J}^*(\sigma_{n+1}, \boldsymbol{v}_{n+1}, \boldsymbol{f}) = \frac{1}{2}(&\|\mathbb{C}_t^{-1/2}\sigma_{n+1} - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}_{n+1}) - \boldsymbol{f}_1\|_{0,\Omega}^2 + \\ &\|(1/\sqrt{k_1})\boldsymbol{v}_{n+1} - \sqrt{k_1}\nabla \cdot \sigma_{n+1} - \boldsymbol{f}_2\|_{0,\Omega}^2 \end{aligned} \tag{5.198}$$

and

$$\begin{aligned} \mathcal{J}_{-1}^*(\sigma_{n+1}, \boldsymbol{v}_{n+1}, \boldsymbol{f}) = \frac{1}{2}(&\|\mathbb{C}_t^{-1/2}\sigma_{n+1} - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v}_{n+1} - \boldsymbol{f}_1\|_{0,\Omega}^2 + \\ &\|(1/\sqrt{k_1})\boldsymbol{v}_{n+1} - \sqrt{k_1}\nabla \cdot \sigma_{n+1} - \boldsymbol{f}_2\|_{-1,\Omega}^2 \end{aligned} \tag{5.199}$$

where $\boldsymbol{f}$, $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are defined as for the functionals $\mathcal{J}$ and $\mathcal{J}_{-1}$. For these functionals the following theorem provides the coercivity estimates.

Theorem 13 *For $k_1 > 0$ there exist positive constants c_1, c_2 independent of λ, such that*

$$\frac{1}{c_1} M(\boldsymbol{w}, \tau) \leq \mathcal{J}^*(\tau, \boldsymbol{w}, 0) \leq c_1 M(\boldsymbol{w}, \tau) \tag{5.200}$$

and

$$\frac{1}{c_2} M_{-1}(\boldsymbol{w}, \tau) \leq \mathcal{J}^*_{-1}(\tau, \boldsymbol{w}, 0) \leq c_2 M_{-1}(\boldsymbol{w}, \tau) \tag{5.201}$$

hold for all $(\boldsymbol{w}, \tau) \in H^1_D(\Omega)^d \times H_N(\mathrm{div}; \Omega)^d$

Proof: The upper bound follows directly from theorem 12. To show the lower bound it will suffice to show that:

$$\mathcal{J}(\tau, \boldsymbol{w}, 0) \leq c_{11} \mathcal{J}^*(\tau, \boldsymbol{w}, 0) \tag{5.202}$$

and

$$\mathcal{J}_{-1}(\tau, \boldsymbol{w}, 0) \leq c_{12} \mathcal{J}^*_{-1}(\tau, \boldsymbol{w}, 0). \tag{5.203}$$

This task reduces to showing that $\|\tau - \tau^T\|^2_{0,\Omega}$ can be bounded from above by $\mathcal{J}^*_{-1}$. A result from [40] shows that:

$$\|\mathbb{C}_t^{-1/2}\tau - \mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|^2_{0,\Omega} \geq \frac{1}{2k_1\mu_s} \left\| \frac{\tau - \tau^T}{2} \right\|^2_{0,\Omega}. \tag{5.204}$$

Hence Eq. (5.202) and Eq. (5.203) hold which completes the proof. □

5.8.3 Coercivity in $\boldsymbol{H^1 \times H(\mathrm{div})}$

Some more results from [37] will be used to establish coercivity in $H^1_D(\Omega)^d \times H(\mathrm{div}; \Omega)^d$. There exists a positive constant c_1 independent of λ such that:

$$\frac{1}{c_1}(\|\boldsymbol{v}\|^2_{1,\Omega} + \lambda\|\nabla \cdot \boldsymbol{v}\|^2_{0,\Omega}) \leq \|\mathbb{C}_t^{1/2}\varepsilon(\boldsymbol{v})\|^2_{0,\Omega} \leq c_1(\|v\|^2_{1,\Omega} + \lambda\|\nabla \cdot \boldsymbol{v}\|^2_{0,\Omega}). \tag{5.205}$$

Furthermore there exist constants c_2 and c_3 such that

$$\frac{1}{c_2}\|\tau\|^2_{0,\Omega} \leq \|\mathbb{C}_t^{-1/2}\tau\|^2_{0,\Omega} + \|\nabla \cdot \tau\|^2_{-1,D,\Omega} \leq c_2\|\tau\|^2_{0,\Omega} \tag{5.206}$$

and

$$\frac{1}{c_3}\|\tau\|^2_{H(\mathrm{div};\Omega)} \leq \|\mathbb{C}_t^{-1/2}\tau\|^2_{0,\Omega} + \|\nabla \cdot \tau\|^2_{0,\Omega} \leq c_3\|\tau\|^2_{H(\mathrm{div};\Omega)} \tag{5.207}$$

The proof of these estimates uses a Helmholtz decomposition and can be found in [37].

Theorem 14 *There exist positive constants c_1, c_2, c_3, c_4 independent of λ, such that*

$$\frac{1}{c_1}(\|\tau\|^2_{H(\mathrm{div};\Omega)} + \|\boldsymbol{w}\|^2_{1,\Omega}+\lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \leq \mathcal{J}(\tau,\boldsymbol{w},0) \leq c_1(\|\tau\|^2_{H(\mathrm{div};\Omega)} + \|\boldsymbol{w}\|^2_{1,\Omega} + \lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \tag{5.208}$$

$$\frac{1}{c_2}(\|\tau\|^2_{H(\mathrm{div};\Omega)} + \|\boldsymbol{w}\|^2_{1,\Omega}+\lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \leq \mathcal{J}^*(\tau,\boldsymbol{w},0) \leq c_2(\|\tau\|^2_{H(\mathrm{div};\Omega)} + \|\boldsymbol{w}\|^2_{1,\Omega} + \lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \tag{5.209}$$

$$\frac{1}{c_3}(\|\tau\|^2_{0,\Omega} + \|\boldsymbol{w}\|^2_{1,\Omega}+\lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \leq \mathcal{J}_{-1}(\tau,\boldsymbol{w},0) \leq c_3(\|\tau\|^2_{0,\Omega} + \|\boldsymbol{w}\|^2_{1,\Omega} + \lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \tag{5.210}$$

and

$$\frac{1}{c_4}(\|\tau\|^2_{0,\Omega} + \|\boldsymbol{w}\|^2_{1,\Omega}+\lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \leq \mathcal{J}^*_{-1}(\tau,\boldsymbol{w},0) \leq c_4(\|\tau\|^2_{0,\Omega} + \|\boldsymbol{w}\|^2_{1,\Omega} + \lambda\|\nabla\cdot\boldsymbol{w}\|^2_{0,\Omega}) \tag{5.211}$$

hold for all $(\boldsymbol{w},\tau) \in H^1_D(\Omega)^d \times H_N(\mathrm{div};\Omega)^d$

Proof: These estimates are a consequence of Eqs. (5.205–5.207) and theorems 12 and 13. □

With these results the analysis of the transient problem is comparatively easy.

5.8.4 Stability for the Transient Problem

Theorem 15 *Let $(\boldsymbol{v}_0,\sigma_0) \in H^1_D(\Omega)^d \times H(\mathrm{div};\Omega)^d$. Then there exists a positive constant c_1 which is independent of N such that the solution obtained by using N times the formulation based on functional Eq. (5.180) satisfies the following energy estimate:*

$$\frac{1}{c_1}(\|\sigma_{N+1}\|^2_{H(\mathrm{div};\Omega)} + \|\boldsymbol{v}_0\|^2_{1,\Omega}+\lambda\|\nabla\cdot\boldsymbol{v}_{N+1}\|^2_{0,\Omega}) \leq c_1(\|\sigma_0\|^2_{H(\mathrm{div};\Omega)} + \|\boldsymbol{v}_0\|^2_{1,\Omega} + \lambda\|\nabla\cdot\boldsymbol{v}_0\|^2_{0,\Omega}) + \Delta t\sum_{n=1}^{N}\|((\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}))^T,(\nabla\cdot\sigma)^T)^T\|_{0,\Omega} \tag{5.212}$$

Proof: The theorem directly follows from theorem 8 and estimate Eq. (5.209). □.

In contrast to the *vm*-formulation, the *vσ*-formulation does not need unnaturally strong boundary conditions and no additional stabilising equations. But the drawback is clearly that it therefore requires a subspace of $H(\mathrm{div};\Omega)$ to approximate the stresses. Nonetheless the direct availability of the stress tensor as a problem unknown is clearly an advantage of this formulation.

5.8.5 Recovery of the Displacements

In the $v\sigma$-formulation the structural displacements are not directly available as problem unknowns. But usually the displacements are of great interest. For the FSI problems they are needed to compute the deformations of the computational grid for the fluid domain. Hence the displacements have to be recovered from the available fields.

One possibility is to use the temporal relation between the velocity field and the displacement field:

$$\frac{\partial \boldsymbol{u}}{\partial t} = \boldsymbol{v} \quad \text{in} \quad \Omega \tag{5.213}$$

Then theoretically an arbitrary time discretisaton could be used to obtain the displacements at each time step. With the goal of high accuracy and stability, the theta θ-method with $\theta = 0.5$ seems to be the method of choice. Implicit methods could be used without any additional computational effort as the velocities are known at timestep n and $n+1$ when the displacements should be recovered. The method would read:

$$\boldsymbol{u}_{n+1} = \boldsymbol{u}_n + \Delta t(\theta \boldsymbol{v}_{n+1} + (1-\theta)\boldsymbol{v}_n) \tag{5.214}$$

Unfortunately, some computational tests have indicated that the displacements which were found by this procedure have a tendency to "drift" away. Although this observation would require a more thorough analysis, it is probably related to the lack of conservation of the least squares FEM.

Alternatively the spatial relation between the displacements and the stress tensor σ:

$$\sigma = \mathbb{C}\varepsilon(\boldsymbol{u}) \quad \text{in} \quad \Omega \tag{5.215}$$

could be used. As this partial differential equation would not yield a unique solution, the displacement field has to be fixed on Γ_D:

$$\boldsymbol{u} = 0 \quad \text{on} \quad \Gamma_D. \tag{5.216}$$

To find a solution of this partial differential equation any possible numerical method could be used. Here again the least squares FEM should be used because this way of recovering the displacements fits well in the overall concept.

The standard L^2 least squares functional is then:

$$\mathcal{J}(\boldsymbol{u}) = \frac{1}{2}\|\sigma - \mathbb{C}\varepsilon(\boldsymbol{u})\|^2_{0,\Omega} \tag{5.217}$$

A detailed analysis of this least squares procedure was not carried out. But in the computational tests it worked quite reliably. It should be noted that the system seems to be overdetermined as there are four equations but only two unknowns.

5.9 Physical Weighting

In this section the weights which must be used to get a dimension independent formulation will be examined. Again all units will be in the SI-System. As these least squares formulations are intended to be coupled with the least squares formulations for the fluid, using the same unit as basis makes sense. Then putting the residuals of both equations together will still be physically meaningful. Only the vH-formulation and the $v\sigma$-formulation will be considered.

For the fluid-part the selected basis unit was the acceleration $\boldsymbol{a}[m/s^2]$. Thus Eq. (5.124) and Eq. (5.174) only have to be divided by ρ_s to have the equation residual in terms of accelerations.

The unit of the residual of Eq. (5.135) and Eq. (5.137) is $[1/s]$ and $[1/sm]$ respectively. This implies that the weight α_1^S has the unit $[m/s]$. A possible choice would be

$$\sqrt{\alpha_1^S} = \frac{L}{\Delta t}\left[\frac{m}{s}\right] \tag{5.218}$$

where L denotes a characteristic length of the structure. For α_2^S the unit is $[m^2/s]$. Therefore a possible choice could be:

$$\sqrt{\alpha_2^S} = h\frac{L}{\Delta t}\left[\frac{m^2}{s}\right] \tag{5.219}$$

where h is the characteristic element length. Including the characteristic element length into the weight lets the stabilising term Eq. (5.137) disappear when $h \to 0$. Wheter including the characteristic element length has a negative effect onto the accuracy or not remains an open question. Alternatively $\alpha_2^S = L^2/\Delta t$ could be used.

The situation for the $v\sigma$-formulation is more complicated due to the square root of the elasticity tensor. As the elasticity tensor $\mathbb{C}$ is constructed from the Lamé constants, it has the same unit $[N/m^2]$. Hence the unit of $\mathbb{C}^{-1}$ is $[m^2/N]$ and for $\mathbb{C}^{-1/2}$ one gets $[m/\sqrt{N}]$. Similarly the unit of $\mathbb{C}^{1/2}$ is $[\sqrt{N}/m]$. Therefore the residual of the constitutive law scaled by $\mathbb{C}^{-1/2}$ (Eq. (5.169)) has the unit $[\sqrt{N}/m]$. In the rate formulation Eq. (5.173) the unit of the residual changes to $[\sqrt{N}/ms]$. The scalings, which were proposed for the numerical method change the unit of the residual in the following way:

$$\mathbb{C}^{-1/2}\frac{\partial \sigma}{\partial t} - \mathbb{C}^{1/2}\frac{\partial \varepsilon}{\partial t} = 0 \quad \left[\frac{\sqrt{N}}{ms}\right] \tag{5.220}$$

$$\mathbb{C}^{-1/2}\frac{(\sigma_{n+1} - \sigma_n)}{\Delta t} - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}) = 0 \quad \left[\frac{\sqrt{N}}{ms}\right] \tag{5.221}$$

$$\mathbb{C}^{-1/2}\frac{(\sigma_{n+1} - \sigma_n)}{\sqrt{\Delta t}} - \sqrt{\Delta t}\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}) = 0 \quad \left[\frac{\sqrt{N}}{m\sqrt{s}}\right] \tag{5.222}$$

From the fact that σ has a unit of $[N/m^2]$ it follows that the divergence of the stresses has a unit of $[N/m^3]$, which implies that the unit of the residual of Eq. (5.174) also has this unit. In the numerical procedure the residual is scaled by $\frac{\sqrt{\Delta t}}{\sqrt{\rho_s}}$ which changes the unit of the residual in the following way:

$$\rho_s \frac{\partial \boldsymbol{v}}{\partial t} - \nabla \cdot \sigma = 0 \quad \left[\frac{N}{m^3}\right] \tag{5.223}$$

$$\rho_s \frac{\boldsymbol{v}_{n+1} - \boldsymbol{v}_n}{\Delta t} - \nabla \cdot \sigma = 0 \quad \left[\frac{N}{m^3}\right] \tag{5.224}$$

$$\rho_s \frac{\boldsymbol{v}_{n+1} - \boldsymbol{v}_n}{\sqrt{\Delta t}} - \sqrt{\Delta t} \nabla \cdot \sigma = 0 \quad \left[\frac{N\sqrt{s}}{m^3}\right] \tag{5.225}$$

$$\frac{\sqrt{\rho_s}}{\sqrt{\Delta t}}(\boldsymbol{v}_{n+1} - \boldsymbol{v}_n) - \frac{\sqrt{\Delta t}}{\sqrt{\rho_s}} \nabla \cdot \sigma = 0 \quad \left[\frac{N\sqrt{s}}{m^{3/2}\sqrt{kg}}\right] \tag{5.226}$$

The presence of the square roots within the units of the residuals makes it advantageous to consider the unit of the squared residuals this time. For the squared residual of the constitutive law this results in a unit of $[N/m^2 s]$ while the squared residual of the conservation of momentum yields $[N^2 s/(m^3 kg)]$ which can be simplified to $[N/m^2 s]$. Therefore the squared residual of both equations has exactly the same unit, which makes any additional scaling unnecessary and leads to the conclusion that $\alpha_0^S = \alpha_1^S = 1$.

Finally the symmetry constraint will be considered. It is a constraint on the stress tensor which has the unit $[N/m^2]$. This leads to a unit of $[N^2/m^4]$ for the squared equation residual. A possible choice for α_2^S would be:

$$\alpha_2^S = \frac{\Delta t}{\rho_s h^2} \left[\frac{ms}{kg}\right] \tag{5.227}$$

which multiplied by the squared residual with unit $[N^2/m^4]$ would give:

$$\left[\frac{N^2}{m^4}\right] \cdot \left[\frac{ms}{kg}\right] = \left[\frac{kgN}{s^2 m^3}\right] \cdot \left[\frac{ms}{kg}\right] = \left[\frac{N}{sm^2}\right] \tag{5.228}$$

which fits well with the other parts of the functional. Alternatively the mesh dependent parameter h^2 could be replaced by the square of a characteristic length $\bar{L}$. As there is no mathematical evidence that the symmetry constraint has to be included into the functional, it seems doubtful if its weight should be increased on refined meshes.

5.10 Numerical Tests

In this section, the quality of the three different least squares formulations for the equations describing a transient linear elasticity problem will be evaluated. The evaluation

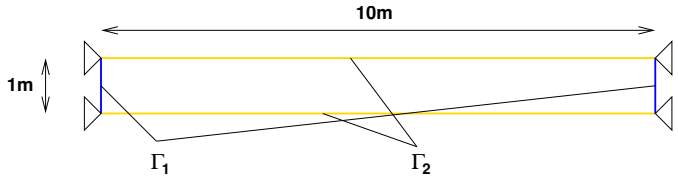

Figure 5.1. Setup of the elastic bar used for the verification

of the results is based on a reference solution. This solution will be computed by using the standard Galerkin formulation, which was shown in section 5.3. To ensure that the implementation of this scheme works correctly, in a preliminary test the results are compared with results found by a commercial finite element code.

After that the test case will be introduced together with the results found by the numerical scheme based on the Galerkin variational principle. Then the same test will be computed using the three different least squares formulations.

5.10.1 Problem Setup and Results for the Galerkin Method

The setup shown in Fig. 5.1 consists of an elastic bar, which is fixed at both ends in all directions (i.e. the displacements or velocities in x- and y-direction are prescribed as Dirichlet boundaries). On the top and bottom of the bar a free boundary condition should be prescribed. For the Galerkin formulation this boundary conditions is the natural boundary condition and therefore does not need any special treatment. For the least squares formulation, the boundary conditions will be discussed in the corresponding section. The material parameters will be uniformly $E = 100{,}000 N/m^2$ and $\nu_s = 0.35$ with a specific mass density of $\rho_s = 1 kg/m^3$.

In the first test a downward acceleration of $1 m/s^2$ is prescribed. The stationary and transient solution of this problem were computed with ANSYS 7.0 using a uniform triangulation of 160 quadrilateral PLANE82 Elements in the *plane strain* configuration (for details see [2]). The maximum static displacement was $u_x = -0.42254 \cdot 10^{-3}$ and $u_y = -0.30935 \cdot 10^{-2}$.

The initial conditions for the transient problem were $u_x = u_y = 0$ in Ω at $t = 0$. As it is well known that structural elements based on a Galerkin variational principle are too stiff, this effect was weakened by using higher order elements. Therefore the model used in the structural code was based on 10 evenly distributed bicubic quadrilateral elements (Q3).

In Fig. 5.2 the vertical displacement u_y at the center node, which is located at (5,1) is plotted over time. The solution found by the newly implemented code with a time step size $\Delta t = 0.001$ is in good agreement with the solution found by ANSYS.

Another test to evaluate the quality of the transient behaviour is the evaluation of the eigenfrequencies, which ANSYS predicted to be $3.253 Hz$, $8.28 Hz$ and $14.879 Hz$.

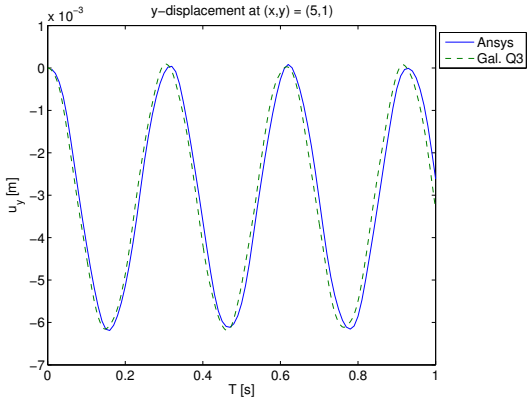

Figure 5.2. Time evolution of u_y at mid node (Galerkin)

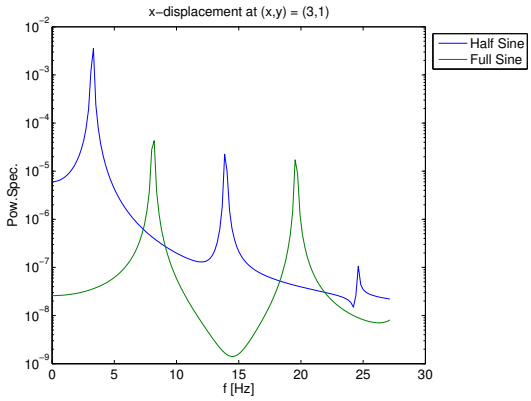

Figure 5.3. Spectrum of displacement u_x at point (3,1) (Galerkin)

ANSYS used a Block-Lanczos method to find the solution of the eigenvalue problem. The new code was tested by prescribing an initial velocity distribution, which was $v_x = 0$, $v_y = \sin(\pi x/10)$ for the first test and $v_x = 0$, $v_y = \sin(\pi x/5)$ for the second test. After that the spectrum of the time series found by computing 512 time steps of size $\Delta t = 0.01$ was analysed. The results shown in Fig. 5.3 indicate that the peaks in the spectrum correspond well with the eigenfrequencies found by ANSYS.

These results demonstrate that the implementation of the Galerkin formulation does work properly. Therefore it can be used as a reliable reference to evaluate the quality of the least squares formulations in the subsequent steps.

For the test of the least squares formulations the same geometry and the same boundary conditions will be used. Again the sinusoidal initial velocity distribution $u_x = 0$, $u_y = \sin(0.1\pi x)$ is prescribed. This initial kinetic energy should slowly be transformed into elastic deformation energy and vice versa. With a fully energy conserving scheme the bar should oscillate indefinitely.

Hence the first important parameter which will be evaluated to judge the quality of the least squares schemes is the total energy of the system throughout time. The kinetic energy is (cf. [74])

$$e_{kin} = \frac{1}{2} \int_\Omega |\boldsymbol{v}|^2 \, d\Omega. \tag{5.229}$$

For the elastic deformation energy of an linear elastic solid, the following holds (cf. [74])

$$e_{def} = \frac{1}{2} \int_\Omega \varepsilon : \mathbb{C}\varepsilon \, d\Omega. \tag{5.230}$$

The results for the Galerkin formulation with 10 elements and $\Delta t = 0.01$ are shown in Fig. 5.4. As it could be expected, the total energy of the system stays constant throughout time.

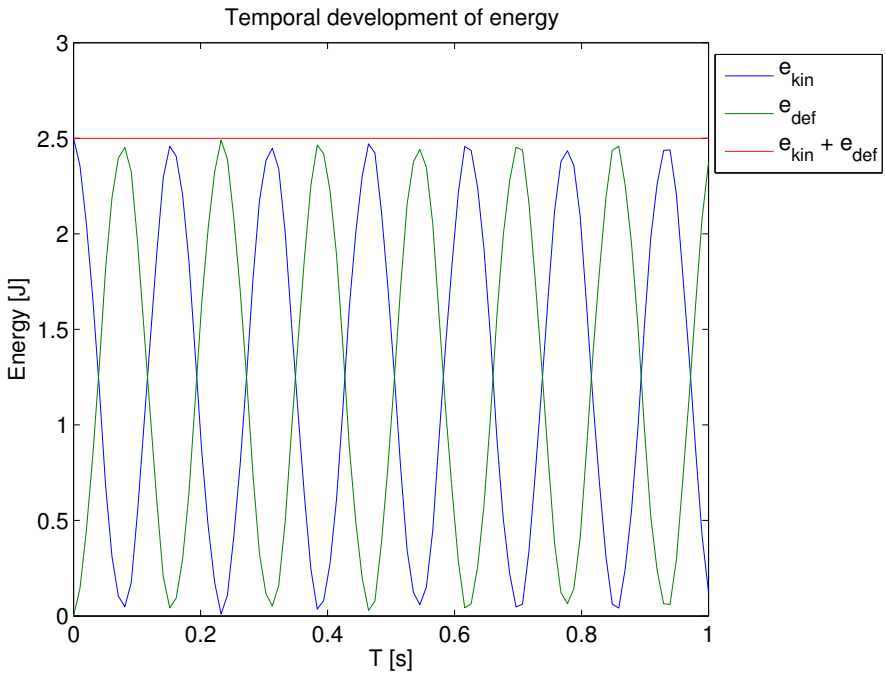

Figure 5.4. Temporal development of the the energy of the elastic bar (Galerkin)

5.10.2 Results of the *vm*-Formulation

Unfortunately, the boundary conditions which are required from the mathematical or numerical side by the vm-formulation do not correspond to the problem related boundary conditions. Actually prescribing only the velocities, which would correspond to a fixed boundary, is not an admissible boundary condition. Therefore the example presented here should only demonstrate that the predictions of theorem 7 could be verified numerically.

Hence the boundary conditions are:

$$v_x = 0 \quad \text{on} \quad \Gamma_1 \tag{5.231}$$
$$\boldsymbol{n} \cdot \mathbf{C}\boldsymbol{m} = 0 \quad \text{on} \quad \Gamma_1, \Gamma_2 \tag{5.232}$$
$$\boldsymbol{n} \times \mathbf{C}\boldsymbol{m} = 0 \quad \text{on} \quad \Gamma_2 \tag{5.233}$$

The initial conditions were set to be:

$$\boldsymbol{m}_0 = 0 \quad \text{in} \quad \Omega \tag{5.234}$$
$$v_{x0} = 0 \quad \text{in} \quad \Omega \tag{5.235}$$
$$v_{y0} = \sin(\pi x/5) \quad \text{in} \quad \Omega \tag{5.236}$$

In Fig. 5.5 the temporal development of the scaled H^1-norm for which stability estimate Eq. (5.101) holds is shown. This problem specific norm decreases slightly with every timestep, which is in compliance with theorem 7 and numerically confirms the stability. During some small time intervals, where the curve becomes horizontal, the method becomes even conservative in the scaled H^1-norm.

But as already explained, the unnatural boundary conditions required by this formulation still prevent this method from being of practical use.

5.10.3 Results of the *vH*-Formulation

For the vH-formulation the following boundary conditions which are compliant with the requirements of the stability estimate were applied:

$$\boldsymbol{v} = 0 \quad \text{on} \quad \Gamma_1 \tag{5.237}$$
$$\boldsymbol{\tau} \cdot \mathbf{H} = 0 \quad \text{on} \quad \Gamma_1 \tag{5.238}$$
$$\boldsymbol{n} \cdot \mathbb{A}\mathbf{H} = 0 \quad \text{on} \quad \Gamma_2 \tag{5.239}$$

The initial conditions were set to:

$$\mathbf{H}_0 = 0 \quad \text{in} \quad \Omega \tag{5.240}$$
$$v_{x0} = 0 \quad \text{in} \quad \Omega \tag{5.241}$$
$$v_{y0} = \sin(\pi x/5) \quad \text{in} \quad \Omega \tag{5.242}$$

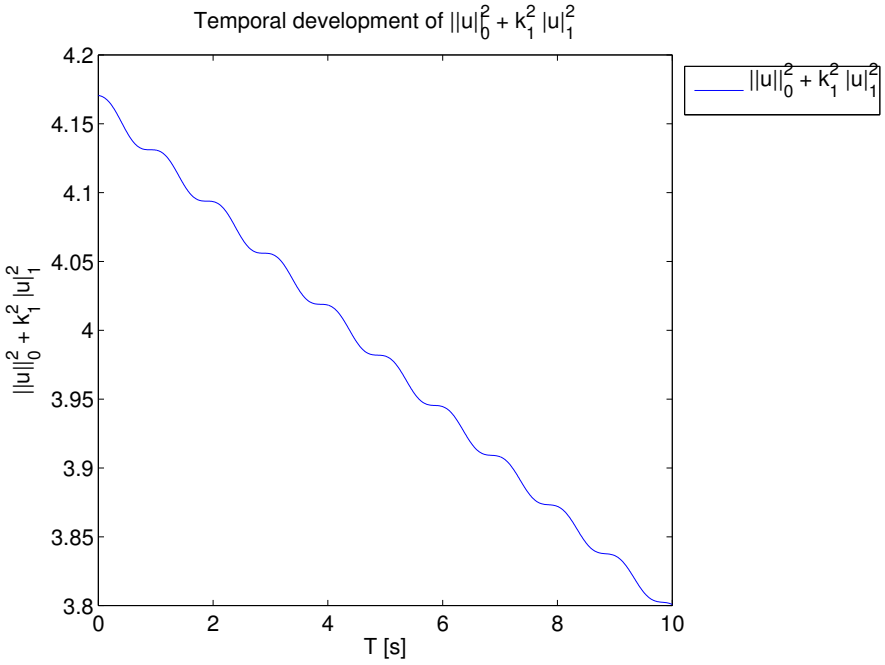

Figure 5.5. Development of H^1-like norm (vm-formulation)

and a time step size of $\Delta t = 0.01$ was selected for all computations. The parameter θ of the time discretisation was set to 0.5 which gives the Crank-Nicholson method. For the weights α_0^S and α_2^S a value of 1.0 was used.

In a first test different finite element meshes were used which were generated from the original 10 element mesh by uniform refinement. Isoparametric bicubic (Q3) elements were used in this test for all unknowns, and the weight α_1^S was set to 1.2, which gave a stable scheme for these parameters. Fig. 5.6 shows the influence of the h-refinement on the temporal development of the energy. These results indicate that the mesh refinement does not have a strong influence onto the energy conservation. On the finest mesh the energy conservation becomes even worse.

To analyse the influence of the weight α_1^S, several computations with different values of α_1^S were performed on a uniform discretisation with 40 Q3-elements. In Fig. 5.7 the results are plotted for $\alpha_1^S = 1.05, 1.1, 1.2$. A clear tendency is that higher values of α_1^S lead to significantly more energy dissipation. But the curve for $\alpha_1^S = 1.05$ shows that the range of possible α_1^S is limited as the formulation becomes slightly unstable if α_1^S is chosen too big. The term *slightly unstable* means in this context that the solution does not blow up within a few timesteps.

The exact reason for this slight instability remains unclear, but is probably related to approximation errors, as it was shown mathematically that the formulation is stable in

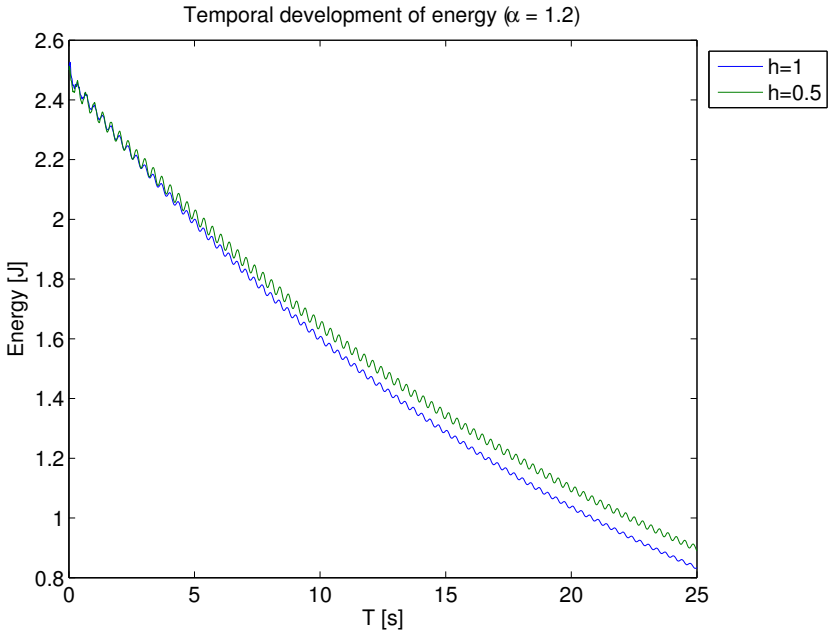

Figure 5.6. Influence of h-refinement onto energy development(vH-formulation)

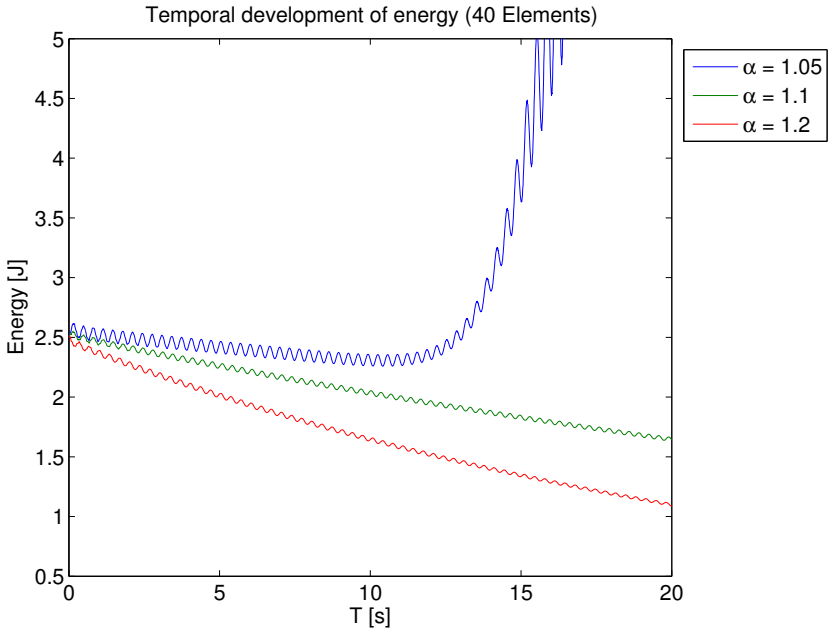

Figure 5.7. Influence of weight α_1^S onto energy development(vH-formulation)

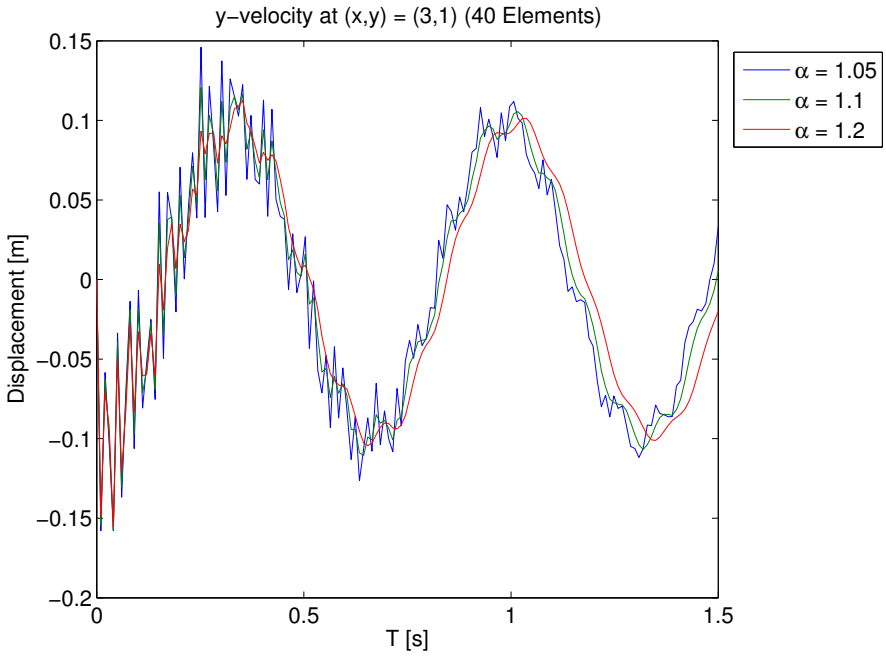

Figure 5.8. Damping of high frequency components in dependence of weight α_1^S

H^1. But the temporal development of v_y which is shown in Fig. 5.8 indicates that the damping of temporal oscillations in the velocity field is influenced by the parameter α_1^S. The curve for $\alpha_1^S = 1.2$ is much smoother at $t = 1.5$ than the one for $\alpha_1^S = 1.05$.

Finally a spectral analysis was performed to get an idea of the accuracy of the vH-formulation . It was carried out on the temporal development of the horizontal velocity v_x, which was found by a computation with $\alpha_1^S = 1.1$ on the grid with 40 elements. The results shown in Fig. 5.9 do not match the results of the Galerkin formulation. All observed frequencies are lower than the ones of the Galerkin formulation. Thus the least squares formulation leads to a softer structure. For higher frequencies the amplitude difference between the results computed with the Galerkin formulation and the ones found with the vH-formulation becomes larger. This leads to the conclusion that higher frequencies will be damped stronger than lower frequencies.

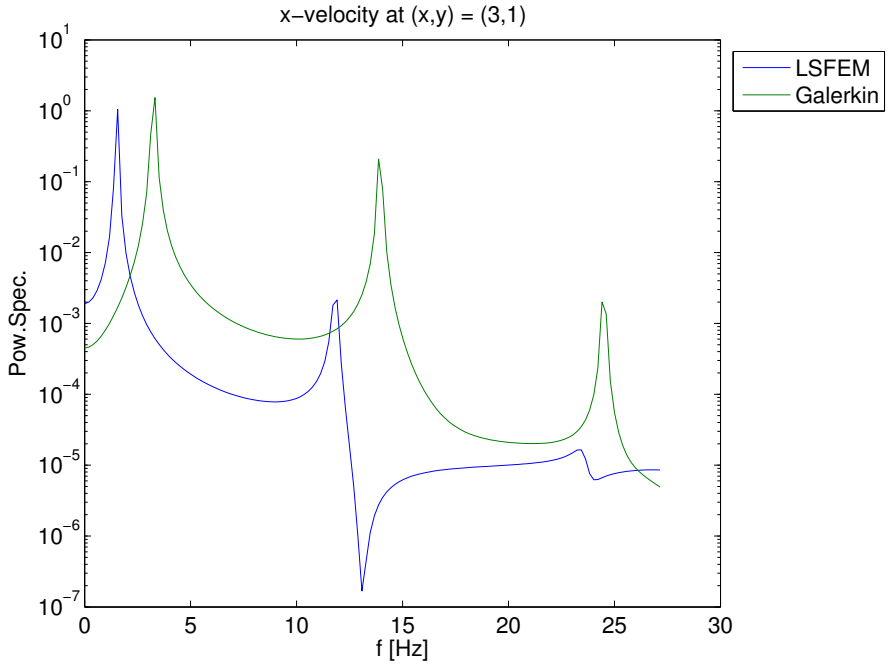

Figure 5.9. Comparison Galerkin/LSFEM in the frequency domain (vH-formulation)

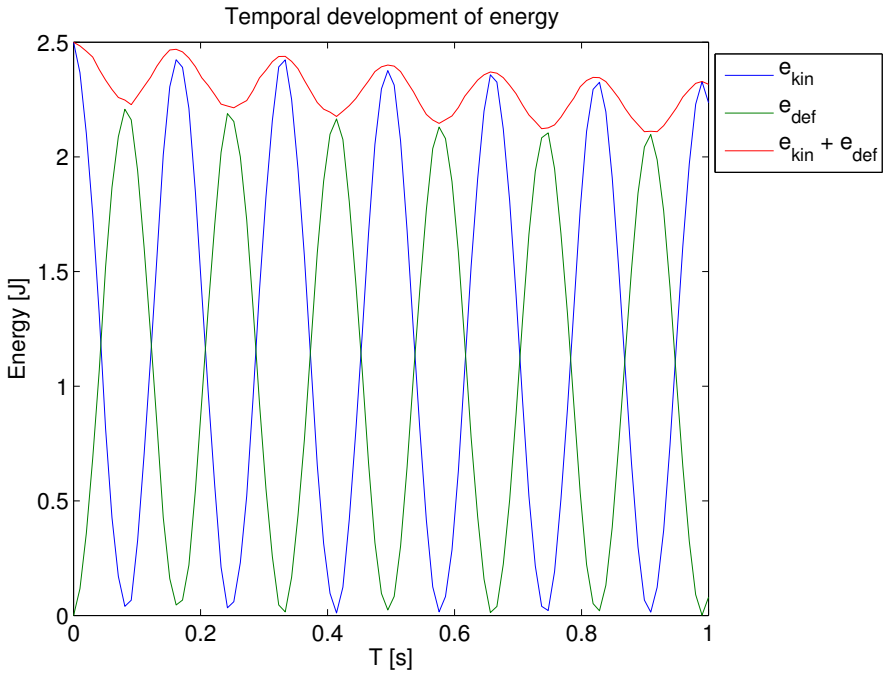

Figure 5.10. Development of energy in elastic bar ($v\sigma$-formulation)

5.10.4 Results of the $v\sigma$-Formulation

For the test of the $v\sigma$-formulation the same setup was used. Similar to the Galerkin test, two different sets of initial conditions were used:

$$\sigma_0 = 0 \quad \text{in} \quad \Omega \tag{5.243}$$
$$v_{x0} = 0 \quad \text{in} \quad \Omega \tag{5.244}$$
$$v_{y0} = \sin(\pi x/5) \quad \text{in} \quad \Omega \tag{5.245}$$

and

$$\sigma_0 = 0 \quad \text{in} \quad \Omega \tag{5.246}$$
$$v_{x0} = 0 \quad \text{in} \quad \Omega \tag{5.247}$$
$$v_{y0} = \sin(\pi x/10) \quad \text{in} \quad \Omega \tag{5.248}$$

This time the different initial conditions were used to analyse the frequency dependence of the numerical damping introduced by the method.

All computations were performed with a time step size $\Delta t = 0.01$ and with element pairs consisting of Q3 elements for the velocity and second order Raviart-Thomas spaces.

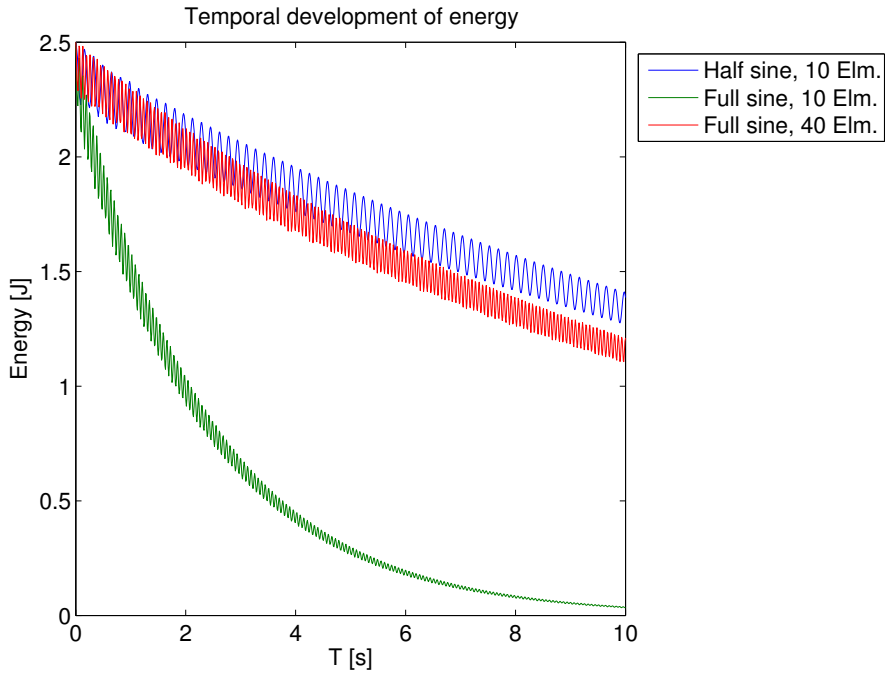

Figure 5.11. Influence of frequency and grid onto energy development($v\sigma$-formulation)

In Fig. 5.10 the temporal development of the kinetic energy, internal deformation energy and total energy is plotted for a time interval of 1 second. Comparing these results with those obtained for the Newmark formulation (cf. Fig. 5.4), the total energy stays not constant over time in the least squares formulation. Instead it oscillates while the energy changes between internal deformation energy and kinetic energy. This can also be seen by looking on both components as the peak of the internal deformation energy is below the peak of the kinetic energy. Hence the transformation between these two forms is not conservative. Nevertheless the overall energy slowly decreases which indicates that the method is stable.

Fig. 5.11 displays the influence of the discretisation parameter h and the initial conditions. With initial conditions Eqs. (5.246–5.248) the energy decreases much faster. Due to the smaller wavelength of these initial conditions, more oscillations in the domain of higher frequencies will be initiated. In this case the numerical test showed that the total energy decreases faster which leads to the conclusion that the higher frequencies will be damped stronger.

Comparing the results of the test with 40 elements and the one with 10 elements shows that the energy conservation becomes much better with decreasing mesh parameter h.

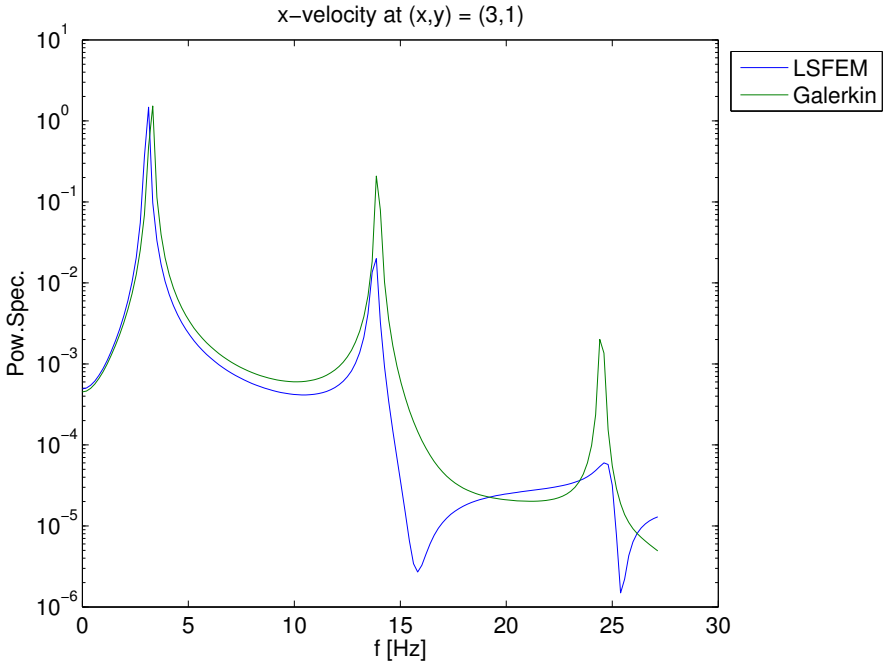

Figure 5.12. Spectrum of velocity v_x at point (3,1) ($v\sigma$-formulation)

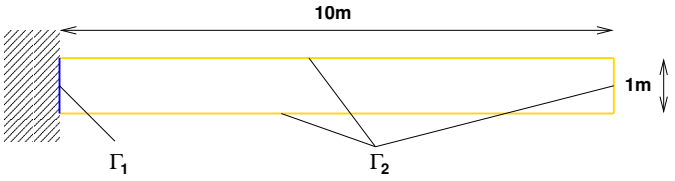

Figure 5.13. Schematic display of the elastic flap

Finally the power spectrum of the time series obtained with the $v\sigma$-formulation was computed and plotted together with the results for the Galerkin formulation in Fig. 5.12. A first property which can be seen in the diagram is again the stronger damping of the higher frequencies because the amplitude gap for the higher eigenfrequencies is larger than that for the lower eigenfrequencies. Furthermore the diagram shows that the results obtained with the $v\sigma$-formulation are in very good agreement with the results computed with the Galerkin formulation.

5.10.5 Influence of the Symmmetry Constraint for the $v\sigma$-Formulation

As it was shown in section 5.8, including the symmetry of the stress σ as additional constraint into the least squares functional is not required from the mathematical point of view. In this section some computational tests, which give an impression of the effects caused by including this constraint or not, will shortly be shown. The first test is related to the strongly coupled example which will be presented in chaper 7. It is the simulation of an elastic flap, which on one end is attached to a rigid wall (cf. Fig. 5.13). The boundary conditions are:

$$\boldsymbol{v} = 0 \quad \text{on} \quad \Gamma_1 \tag{5.249}$$

$$\sigma \cdot \boldsymbol{n} = 0 \quad \text{on} \quad \Gamma_2 \tag{5.250}$$

and the material parameters are $E = 2.5 \cdot 10^6 N/m^2, \nu_s = 0.35, \rho_s = 0.1 kg/m^3$ To trigger the deformation, an uniform acceleration $\boldsymbol{a} = (0, -1)^T m/s$ is applied as right hand side term.

The stationary solution of that problem was computed by using a formulation based on Eqs. (5.169–5.170). Clearly the transient solution should settle at this solution when $t \to \infty$. Two element pairs, $RT2 - Q2$ and $RT2 - Q3$, were tested on uniformly refined meshes. The weights were chosen to be $\alpha_0^S = \alpha_1^S = \alpha_3^S = 1$. First no symmetry constraint was added ($\alpha_2^S = 0$). After that the computations were repeated with the symmetry constraint ($\alpha_2^S = 1$). In all tests Eq. (5.250) was implemented in a least squares sense and weighted with 1.

Table 5.1 summarises results and sets them also in relation to a solution found with ANSYS using a discretisation with 1000 elements of type PLANE82. The difference between the LSFEM solution and the ANSYS solution remains unclear. But the results

h	RT2-Q2/Sym	RT2-Q3/Sym	RT2-Q2/No Sym	RT2-Q3/No Sym
2^0	$5.864 \cdot 10^{-4}$	$5.91966 \cdot 10^{-4}$	$0.88419 \cdot 10^{-4}$	$1.87165 \cdot 10^{-4}$
2^{-1}	$5.92698 \cdot 10^{-4}$	$5.94235 \cdot 10^{-4}$	$2.59164 \cdot 10^{-4}$	$3.77642 \cdot 10^{-4}$
2^{-2}	$5.95026 \cdot 10^{-4}$	$5.95534 \cdot 10^{-4}$	$4.36456 \cdot 10^{-4}$	$4.99988 \cdot 10^{-4}$
2^{-3}	$5.95919 \cdot 10^{-4}$	$5.96104 \cdot 10^{-4}$	$5.30089 \cdot 10^{-4}$	$5.57573 \cdot 10^{-4}$

Table 5.1. Maximum vertical displacement of elastic flap computed with the $v\sigma$-formulation for linear elasticity (Ansys solution: $5.31 \cdot 10^{-4}$).

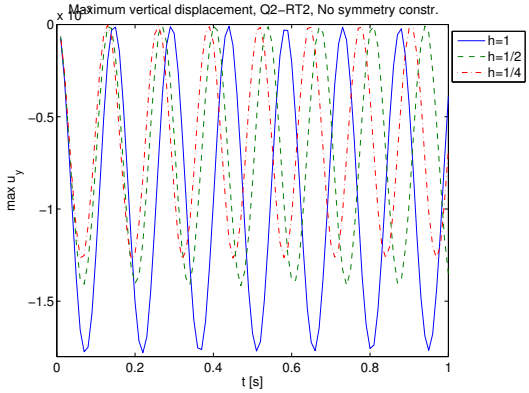

Figure 5.14. Time evolution of $\max u_y$, without symmetry constraint

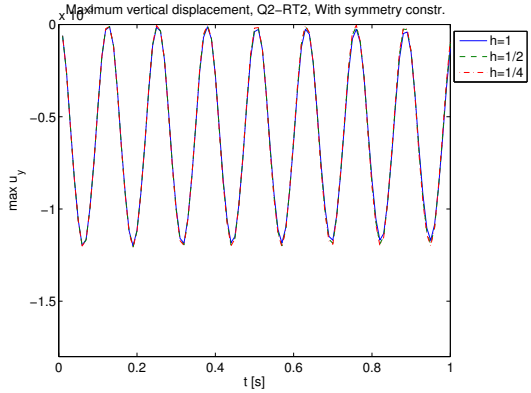

Figure 5.15. Time evolution of $\max u_y$, with symmetry constraint

in Table 5.1 indicate that the least squares solutions converge to a maximum vertical displacement of about $5.96 \cdot 10^{-4}$. Thus the formulation without the symmetry constraint achieves much lower accuracy on a given discretisation for this test case.

To evaluate the effect of the symmetry constraint onto the $v\sigma$-formulation for transient problems, the time dependent solution was computed with and without the symmetry constraint using $RT2 - Q2$ element pairs and a time step size of $\Delta t = 0.01$.

The results shown in Fig. 5.14 and Fig. 5.15 show a very interesting effect. While not sufficiently fine discretisations usually lead to an underestimation of the displacements in the Galerkin context, they lead to an overestimation of the displacements in the $v\sigma$-formulation! If the symmetry constraint is included, the quality of the approximation seems to be much better. Even on coarse discretisations, the results are reasonable.

These results motivated another computational examination to find out whether this effect can also be seen on more synthetic convergence tests. For this purpose a solution field:

$$\boldsymbol{u} = (\sin(2\pi x)(y - y^2), \sin(2\pi y)(x - x^2))^T \quad \text{in} \quad \Omega \tag{5.251}$$

with boundary conditions

$$\boldsymbol{u} = 0 \quad \text{on} \quad \Gamma \tag{5.252}$$

h	$\|u_x - u_{h_x}\|_{0,\Omega}$	Rate	$\|\sigma_{11} - \sigma_{h_{11}}\|_{0,\Omega}$	Rate	$\|\sigma_{12} - \sigma_{h_{12}}\|_{0,\Omega}$	Rate
2^0	$1.291912 \cdot 10^{-1}$	—	$1.051064 \cdot 10^6$	—	$3.952869 \cdot 10^5$	—
2^{-1}	$5.158451 \cdot 10^{-3}$	25.045	$1.066440 \cdot 10^5$	9.856	$1.851969 \cdot 10^4$	21.344
2^{-2}	$2.815793 \cdot 10^{-3}$	1.832	$2.957591 \cdot 10^4$	3.606	$9.115512 \cdot 10^3$	2.032
2^{-3}	$3.584496 \cdot 10^{-4}$	7.855	$7.115234 \cdot 10^3$	4.157	$1.395716 \cdot 10^3$	6.531
2^{-4}	$4.492952 \cdot 10^{-5}$	7.978	$1.765259 \cdot 10^3$	4.031	$1.866447 \cdot 10^2$	7.478
2^{-5}	$5.619056 \cdot 10^{-6}$	7.996	$4.401978 \cdot 10^2$	4.01	$2.397491 \cdot 10^1$	7.785

Table 5.2. Convergence of $v\sigma$-formulation with symmetry constraint

h	$\|u_x - u_{h_x}\|_{0,\Omega}$	Rate	$\|\sigma_{11} - \sigma_{h_{11}}\|_{0,\Omega}$	Rate	$\|\sigma_{12} - \sigma_{h_{12}}\|_{0,\Omega}$	Rate
2^0	$1.291912 \cdot 10^{-1}$	—	$1.016915 \cdot 10^6$	—	$3.235552 \cdot 10^5$	—
2^{-1}	$5.196130 \cdot 10^{-3}$	24.863	$1.067536 \cdot 10^5$	9.526	$2.258909 \cdot 10^4$	14.324
2^{-2}	$3.007495 \cdot 10^{-3}$	1.728	$1.595264 \cdot 10^4$	6.692	$2.253571 \cdot 10^4$	1.002
2^{-3}	$3.685047 \cdot 10^{-4}$	8.161	$1.784453 \cdot 10^3$	8.94	$4.283167 \cdot 10^3$	5.261
2^{-4}	$4.530965 \cdot 10^{-5}$	8.133	$1.969592 \cdot 10^2$	9.06	$6.276504 \cdot 10^2$	6.824
2^{-5}	$5.632072 \cdot 10^{-6}$	8.045	$2.918788 \cdot 10^1$	6.748	$9.193436 \cdot 10^1$	6.827

Table 5.3. Convergence of $v\sigma$-formulation without symmetry constraint

was considered on the unit square which was discretised using $RT2 - Q2$ element pairs. Inserting this solution into Eq. (2.45) gives a right hand side forcing term, which was applied.

The L^2 errors for $u_x, \sigma_{11}, \sigma_{12}$ are shown in Table 5.2 for the formulation with symmetry constraint and in Table 5.3 for the formulation without symmetry constraint respectively. Both formulations achieve optimal convergence rates in the displacement field. In the stress field, the situation is different. While the constrained formulation seems to loose one order of convergence in the unknowns σ_{11} and σ_{22}, it achieves nearly full accuracy for σ_{12}. In the unconstrained formulation the situation is different. Better convergence is achieved in σ_{11} and σ_{22} for the price of lower convergence in σ_{12}.

Unfortunately, these results do not give a clear picture of the situation. While for the elastic flap the computational results indicate that the symmetry constraint is absolutely necessary, it does not seem to have that importance in the synthetic benchmark problem on the unit square. Hence further investigations clearly are required.

5.11 Summary

Interestingly, the available literature about least squares methods for transient problems is very rare. A few articles consider the treatment of convection problems and the Navier-Stokes for instationary flows (cf. [59], [33], [121]). But the mathematical analysis

of these methods is very limited. The space-time least squares FEM was considered in [93] and [94]. For the transport equation some results were presented in [50].

Therefore some more general ideas were developed in the first sections of this chapter. These results were applied to the wave equation and it was demonstrated that in more than 1 dimension the least squares FEM is not fully H^1 stable anymore. Therefore finite elements which are subspaces of the $H(\mathrm{div};\Omega)$ have to be used to get a working numerical method in more than 1 dimension.

Another way to get a stable numerical method is to augment or stabilise the least squares functional by additional equations, which restrict the function space for the solution. For the 2D case this kind of formulation was developed together with an appropriate stability estimate in H^1 for all unknowns. But these seemingly better properties do not come for free. On the one hand the restricted solution space might be too small, on the other hand the formulation requires stronger boundary conditions.

This dilemma became clear with the vm-formulation for the equations of linear elastodynamics. Although the formulation was stable in H^1, physically meaningful boundary conditions were not admitted and hence rendered the method useless for practical applications.

Looking at the principal part of the above mentioned equations, it was always symmetric, which allowed the use of partial integration to prove the stability estimates. Clearly this is a very strong necessity which prevents the analysis of more general equations. This lead to the development of a theorem which shows a slightly weaker stability estimate under less restrictive assumptions.

With this theorem the vH-formulation and the $v\sigma$-formulation were analysed with respect to their stability. Actually these formulations are modified versions of the formulations for the linear elasticity. Using the theory of elliptic systems, full H^1 coercivity was shown for the augmented vH-formulation for elastodynamics with two physically meaningful boundary conditions. It was also demonstrated, that the vH-formulation is not H^1-coercive if it is not augmented by some stabilising terms.

For the $v\sigma$-formulation prior results which were published in [40], [38] and [37] were modified such that they fit the $v\sigma$-formulation for elastodynamics. As this formulation does not use stabilising terms, the coercivity and thus also the stability of the formulation for the transient problem can only be shown in the weaker $H(\mathrm{div};\Omega)$-norm. This behaviour is comparable to the behaviour observed for the wave equation in more than one dimension. The necessary boundary conditions are in compliance with the physically meaningful boundary conditions namely the velocity boundary conditions or the normal tractions.

Clearly all theoretical results have to be verified numerically. This was done in the last section for the Galerkin formulation, the vm-formulation , the augmented vH-formulation and the $v\sigma$-formulation. As the Galerkin formulation is known to work well, the primary objective of that test was to exclude the possibility of implementational errors. The test of the vm-formulation was mainly presented for sake of completeness as

already the mathematical analysis revealed that it will not admit physically meaningful boundary conditions.

The results obtained with the vH-formulation are quite interesting because they demonstrate that the formulation can exhibit a slight numerically instability if the least squares terms are not weighted correctly. What the correct weights are in this context remains an open question. Nonetheless the results are not very convincing. Within the stable regime of the weighting parameter, the formulation becomes quite dissipative and even the eigenfrequencies of the bar are not in good agreement with the solution found by the Newmark-Galerkin scheme.

For the $v\sigma$-formulation the results are much more promising. During the numerical tests instabilities could only be found if the weight of the least squares term which implemented the normal traction boundary condition was chosen extremely high or low. Additionally the results were in good agreement with the solution found by the Newmark-Galerkin scheme.

To conclude the numerical tests, the most promising formulation is the $v\sigma$-formulation which numerically seems to be quite robust and has the advantage of directly providing the real stresses inside the structure. The clear disadvantage is the need for finite element subspaces of $H(\mathrm{div};\Omega)$, namely the Raviart-Thomas spaces, which are a bit laborious with respect to their handling.

Although several new results were found for least squares methods for second order hyperbolic systems and the equations of linear elastodynamics, several questions remain open. The very important aspect of error estimates was completely left out. It might be possible that an error analysis could also provide some insight into reasons for the conditional instability of the vH-formulation . Furthermore a generalisation of the results to negative norm methods could lead to alternative first order formulations for this problem class. Currently all estimates implied a dissipative behaviour. Identifying the corresponding terms in the resulting equations and modifying the right hand side terms accordingly could perhaps restore conservative behaviour and would clearly be an interesting option.

6 Implementation of the Coupling Conditions

In this chapter the numerical implementation of the coupling conditions will be introduced. After a short summary of the coupling conditions in the problem unknowns, the numerical issues will be discussed. As both the LSFEM and the Galerkin formulation will be used for the structural part, the different implementation for these variants will be discussed. Finally some simple numerical tests will be performed to compare the quality of the LSFEM coupling procedure with the quality of the more standard coupling between the Galerkin structural formulation and the LSFEM formulation for the fluid part.

6.1 Numerical Methods for the Coupling Conditions

In various areas of scientific computing the problem of interface conditions comes up. Besides the area of multiphysics, domain decomposition methods are a very important field of application. In multiphysics the interface conditions usually are connected to the physical model. Domain decomposition methods are mainly used to efficiently solve large systems of equations on parallel computers (cf. [110]). For this purpose the domain of the original problem is divided into smaller subdomains. On the interfaces between these subdomains the correct boundary conditions have to be prescribed. Actually it came out that these algorithms have many aspects in common with partitioned algorithms for multiphysics problems (cf. chapter 8 in [110]), which could be considered as a domain decomposition of the strongly coupled problem into two subdomains.

If the interfaces between the subdomains are conforming, that means the finite element functions could be pointwise equivalent, the correct algorithmic treatment is mostly clear. The difficulties begin with nonconforming interfaces where the spatial discretisation on one side of the interface significantly differs from the spatial discretisation on the other side of the interface (cf. p.4 in [131]). Probably the simplest algorithm to treat these nonconforming interfaces is the interpolation of the solution from the boundary of one subdomain onto the boundary of the neighbouring subdomain. Despite its intrigu-

ing simplicity, in its basic form this method has disadvantages regarding stability and conservation properties (cf. [46]).

To overcome these difficulties, other methods were proposed. Three important techniques are the three-field method (cf. [29], section 1.7 in [109]), the mortar element technique (cf. [131]) and the finite element tearing and interpolating (FETI) method (cf. [65]).

Clearly imposing the coupling conditions in the least squares sense is the natural choice when already the fluid and the structural part are formulated with this concept. This method has been proposed for a boundary/finite element coupling in [92]. For FSI problems this approach was suggested in [76],[75] and [54]. Furthermore in [117] some general issues of imposing the boundary conditions only in the weak sense for the least squares FEM have been analysed.

6.2 Coupling Conditions in the Problem Variables

The velocities appear in all proposed formulations for the fluid and the structure as problem unknowns. Hence Eq. (2.49) can be directly used. For the traction coupling condition

$$(\sigma_f - \sigma_s)\,\boldsymbol{n}_C = 0 \quad \text{on} \quad \Gamma_{if} \tag{6.1}$$

the stresses must be expressed in the problem unknowns. As the structure is usually formulated in a Lagrangian description, the Cauchy stress tensor is not directly available. Instead it has to be expressed in terms of the 2nd Piola Kirchhoff stress tensor

$$\sigma_s = \frac{1}{\det \mathbf{F}} \mathbf{F}\mathbf{S}\mathbf{F}^T \tag{6.2}$$

When large deformations should be treated physically correctly, this nonlinear coupling term is obligatory.

In this work, the focus is put onto the least squares variational principle. Therefore the complications induced by the nonlinearity coming from the physically correct treatment of large deformations will be excluded.

The simplified Hooke model which is the basis for the equations of linear elasticity includes the Cauchy stress tensor, which can directly be used in Eq. (6.1). As long as the structural deformations are small, this model is correct. When the deformations become large, the model will be erroneous but sufficient for the objectives of this work, namely the analysis of the principal usability of least squares methods for FSI problems.

Thus the following linear coupling conditions will be used:

$$(-\mathbf{I}p + 2\nu_f \mathbf{D} - (\lambda(\operatorname{tr}\varepsilon(\boldsymbol{u}))\mathbf{I} + 2\mu_s\varepsilon(\boldsymbol{u})))\,\boldsymbol{n}_C = 0 \quad \text{on} \quad \Gamma_{if} \tag{6.3}$$

For the $v\sigma$-formulation of linear elastodynamics, the stress tensor σ_s is directly available as unknown, which simplifies Eq. (6.3) to:

$$(-\mathbf{I}p + 2\nu_f \mathbf{D} - \sigma_s)\,\boldsymbol{n}_C = 0 \quad \text{on} \quad \Gamma_{if} \tag{6.4}$$

6.3 Function Spaces

To properly define the variational formulation of the coupling conditions, a brief discussion of the functional setting of the coupled problem is necessary. For the Navier-Stokes equations for incompressible fluids coupled with the equations of linear elastodynamics, the results are rare and incomplete even in absence of mesh deformations which further complicate the situation (cf. [61]). Thus the discussion will be limited to the Stokes equations for transient flows coupled with the linear equations of elastodynamics in a fixed domain. First some results from [62] and [61] will be presented, which show the existence and uniqueness of weak solutions in the appropriate spaces for the linear coupled problem. After that some remarks about the spaces for the least squares formulation will follow.

6.3.1 Weak Solutions

In the following paragraphs, the spaces will be discussed in which weak solutions of the linear coupled problem exist. Usually these spaces are also the ones which appear within the Galerkin framework for that problem. Most existence and uniqueness results for partial differential equations also rely on the technique of formulating the problem in a weak setting and then analysing the resulting integral forms.

While many linear partial differential equations have been analysed by this technique, only a few results are available for linear coupled problems. For nonlinear coupled problems, the situation is even worse. Unfortunately, the existence and uniqueness of solutions for the subproblems does not imply the existence and uniqueness of solutions for the coupled problem. Therefore the results for the linear equations of elastodynamics (cf. chapter XVIII, section 6 in [57] for example) and the Navier-Stokes equations for incompressible flows (cf. [123], [70], [116]) are of limited use for the coupled problem.

Therefore the discussion of the relevant spaces will mainly rely on the results from [61] and [62], which provide existence and uniqueness results for a linear coupled problem which consists of the Stokes equations for transient flows and the linear equations of elastodynamics. The system of equations, which is considered in [61] and [62] consists of Eq. (2.44) in a domain Ω_1 and Eq. (2.40) without convective term (sometimes denoted by $\eta_f = \infty$) in a domain Ω_2. In the next paragraphs the right hand side terms of Eq. (2.44) and Eq. (2.40) will be denoted y $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ respectively. Both domains share a part of their boundary, which is denoted by Γ_{if}. On the rest of the boundary, the boundary condition is $\boldsymbol{v}_f = 0$ or $\boldsymbol{v}_s = 0$ respectively. The boundary of the both domains is assumed to be Lipschitz continuous. Along Γ_{if} the coupling conditions Eq. (6.3) and Eq. (2.49) should hold. Within the theorems, the following spaces which are the restrictions of H_0^1 onto the subdomains Ω_1 and Ω_2 are required:

$$\mathbf{X}_i = [H_0^1(\Omega)]|_{\Omega_i}, \quad i = 1,2. \tag{6.5}$$

They are equipped with the following norm:

$$\|\cdot\|_{X_i} = \|\cdot\|_{1,\Omega_i}. \tag{6.6}$$

For this coupled system the following theorem, which is a reduced version of theorem 2.2 from [62], shows the existence of a weak solution:

Theorem 16 *Assume that* $\boldsymbol{f}_1 \in C([0,T];L^2(\Omega_1))$, $\boldsymbol{f}_2 \in C([0,T];L^2(\Omega_2))$, $\boldsymbol{u}_0 \in X_2$, $\boldsymbol{v}_{f_0} \in X_1$, $\nabla \cdot \boldsymbol{v}_0 = 0\, in \Omega_1$, $\boldsymbol{v}_{s_0} \in X_2$ *and* $\boldsymbol{v}_{f_0} = \boldsymbol{v}_{s_0}\, on \Gamma_{if}$. *Then, there exists a unique solution* $(\boldsymbol{v}_f, \boldsymbol{u}) \in L^2(0,T;X_1) \times L^2(0,T;X_2)$.

Imposing slightly stronger regularity demands onto the initial conditions and the right hand side terms, the following theorem which is a reduced version of theorem 2.3 from [62] shows the existence of a square integrable pressure:

Theorem 17 *Assume that* $\boldsymbol{f}_1 \in C([0,T];L^2(\Omega_1))$, $\boldsymbol{f}_2 \in C([0,T];L^2(\Omega_2))$, $\boldsymbol{u}_0 \in X_2$, $\boldsymbol{v}_{f_0} \in X_1$, $\nabla \cdot \boldsymbol{v}_0 = 0\, in \Omega_1$, $\boldsymbol{v}_{s_0} \in X_2$, $\boldsymbol{v}_{f_0} = \boldsymbol{v}_{s_0}\, on \Gamma_{if}$, $\frac{\partial}{\partial t}\boldsymbol{f}_i \in L^2(0,T;L^2(\Omega_i)), i = 1,2$, $\boldsymbol{v}_{f_0} \in H^2(\Omega_1)$, $\boldsymbol{v}_{s_0} \in H^1(\Omega_2)$ *and* $\boldsymbol{u}_0 \in H^2(\Omega_2)$. *Assume further that there exists a* $p_0 \in H^1(\Omega_1)$ *such that Eq. (6.3) is satisfied for the initial conditions. Then the solution satisfies:*

$$\boldsymbol{v}_f \in L_\infty(0,T;L^2(\Omega_1)) \cap L^2(0,T;X_1), \quad \boldsymbol{u} \in L_\infty(0,T;X_2) \tag{6.7}$$

$$\frac{\partial \boldsymbol{v}_f}{\partial t} \in L_\infty(0,T;L^2(\Omega_1)) \cap L^2(0,T;X_1), \quad \boldsymbol{v}_s \in L_\infty(0,T;X_2), \quad \frac{\partial \boldsymbol{v}_s}{\partial t} \in L_\infty(0,T;L^2(\Omega_2)) \tag{6.8}$$

Furthermore there exists a unique $p \in L^2(0,T;L^2(\Omega_1))$.

6.3.2 Spaces for the Least Squares Formulations

As it was already mentioned in subsection 3.5.2, many least squares formulations lead to increased regularity demands on the solution spaces. Another point which influences the regularity demands on the solution spaces is the type of the finite dimensional approximation spaces. The benefits of p-refinement are partially lost when the solution does not possess sufficient smoothness (cf. section 3.3). Therefore in the next paragraphs the basic properties and the essential requirements of the solution spaces for the least squares formulation will be discussed with respect to the coupling conditions.

Actually a full analysis would require stability estimates for the least squares functional representing the coupled problem. This task seems to be possible but will not be part of this work. Therefore the discussion will mainly rely on some heuristic reasoning.

Another detail which should be mentioned is the fact that in contrast to the previously shown existence theorems, the applied least squares method includes the temporal discretisation. Thus the solution will only be considered at certain time instances.

Structural Domain

Referring to the results of chapter 5, the appropriate spaces for the $v\sigma$-formulation are:

$$\boldsymbol{v}_s \in H^1(\Omega_s), \quad \varepsilon(\boldsymbol{v}_s) \in L^2(\Omega_s), \quad \sigma_s \in H(\mathrm{div};\Omega_s) \tag{6.9}$$

where Ω_s denotes the structural domain.

Fluid Domain

For the negative norm variants of the fluid formulation, the solutions are sought in the following spaces (cf. [34], [14], [20]):

$$\boldsymbol{v} \in H^1(\Omega_f)^d, \quad p \in L^2(\Omega_f), \quad \omega \in L^2(\Omega_f)^{2d-3}, \quad \mathbf{L} \in L^2(\Omega_f)^{d^2}. \tag{6.10}$$

with Ω_f being the fluid domain. Together with the weighted L^2 formulation, these spaces are too large to achieve optimal convergence rates (cf. [21]). To achieve optimal convergence rates with the weighted $up\omega$-formulation, the minimum regularity requirement is $\boldsymbol{v} \in H^2(\Omega_f)$, $p \in H^1(\Omega_f)$ and $\omega \in H^1(\Omega_f)$.

Interface

For the definition of the interface spaces the functions within the domains Ω_s and Ω_f are assumed to be in the spaces as specified by Eq. (6.9) and Eq. (6.10).

The spaces on the structural side of the interface will be considered first. Following theorem 2.5 in [70], there exists a mapping $\gamma_n : \boldsymbol{v} \to \boldsymbol{v}\cdot\boldsymbol{n}|_\Gamma$ from $H(\mathrm{div};\Omega)$ to $H^{-1/2}(\Gamma)$. This mapping is denoted by $\boldsymbol{v}\cdot\boldsymbol{n}$. Hence a function defining the normal stresses $\sigma_s\cdot\boldsymbol{n}$ on the interface can be an element of $H^{-1/2}(\Gamma_{if})$. The velocity field $\boldsymbol{v}_s$ of the structure is an element of $H^1(\Omega_s)$. Following theorem 9.15 in [105] there exists a bounded linear trace operator:

$$\mathcal{T} : H^k(\Omega) \to H^{k-1/2}(\Gamma) \tag{6.11}$$

for an integer $k > 0$. Hence the restriction of $\boldsymbol{v}_s \in H^1(\Omega_s)$ onto Γ_{if} is an element of $H^{1/2}(\Gamma_{if})$. A function prescribing the velocity field of the structure along the interface could therefore be in $H^{1/2}(\Gamma_{if})$.

The velocity space of the fluid is at least $H^1(\Omega_f)$. Along the interface the velocity will therefore be in $H^{1/2}(\Gamma_{if})$. Minimising the $H^{1/2}$ norm of the residual of Eq. (2.49) should thus be appropriate to include the compatibility condition into the least squares functional.

For the traction interface condition the situation is more complicated. While theorem 2.5 in [70] guarantees that the normal stresses on the structural side of the interface are in $H^{-1/2}(\Gamma_{if})$, no such result is available for the fluid side because the quantities p and $\nabla\boldsymbol{v}_f$ which define the normal stress are only in L^2 with the consequence that their traces on the interface do not even have to exist (cf. section 5.2 in [32]).

Looking at the Galerkin formulation of the Stokes equations:

$$\int_\Omega \nabla \boldsymbol{v} \cdot \boldsymbol{\varphi}\, d\Omega - \int_\Omega p \nabla \varphi\, d\Omega - \int_\Gamma (\boldsymbol{n} \cdot \nabla u)\varphi + \boldsymbol{n} \cdot \varphi p\, d\Gamma = 0 \tag{6.12}$$

it comes out that the normal tractions along the boundary can be prescribed by functions from $H^{-1/2}$. Hence it will be assumed that the normal stresses of the fluid side will be in $H^{-1/2}(\Gamma_{if})$. A detailed analysis of the least squares functional would be required to ensure the validity of this assumption. Some short calculations based on the proofs in [34] and [37] were quite promising, but will not be pursued further here.

6.4 LSFEM-Galerkin Coupling

For Eq. (2.44) the Galerkin variational principle does not correspond to a minimisation problem due to the hyperbolic nature of the equations. Thus, the least squares functional cannot easily be extended with terms representing the structural part and the coupling conditions to get a unified approach.

The velocities in the structure should act on the fluid part by defining the velocity on the interface. This condition is implemented in a weak sense. On the fluid structure interface we demand that the integral of Eq. (2.49) multiplied with some test functions vanishes:

$$\int_{\Gamma_{if}} (\boldsymbol{v}_f - \boldsymbol{v}_s)\varphi\, d\Gamma_{if} = 0. \qquad \forall \varphi \in H^{1/2}(\Gamma_{if}) \tag{6.13}$$

It is not possible to simply add this condition to the bilinear forms stemming from the least squares formulation. Instead it must be implemented as a kind of restriction of the solution space for the fluid unknowns:

$$V_h = \left\{ \boldsymbol{v_f} \in H^1(\Omega) | \int_{\Gamma_{if}} (\boldsymbol{v}_f - \boldsymbol{v}_s)\varphi\, d\Gamma_{if} = 0, \quad \forall \varphi \in H^{1/2}(\Gamma_{if}) \right\} \tag{6.14}$$

The fluid forces act on the structure through the normal stresses on the boundary, which appear as natural boundary conditions in the Galerkin formulation. In particular, the usual Galerkin formulation of Eq. (2.44) becomes:

$$\begin{aligned} \int_{\Omega_s} \rho_s \frac{\partial \mathbf{v}_s}{\partial t} \varphi\, d\,\Omega_s + \int_{\Omega_s} \sigma_s : \nabla \varphi\, d\,\Omega_s + \\ \underbrace{\int_{\Gamma_s} \sigma_s n \cdot \varphi\, d\,\Gamma_s}_{T} = \int_{\Omega_s} f_s \varphi\, d\,\Omega_s, \quad \forall \varphi \in H^1(\Omega_s) \end{aligned} \tag{6.15}$$

From Eq. (2.46) it is clear that along the interface the stress vector in the structure must be opposite equal to the stress vector in the fluid. Because in the current examples

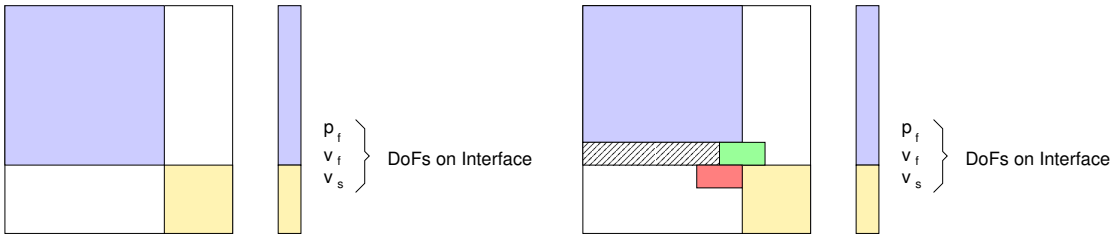

Figure 6.1. Schematic matrix structure for the uncoupled problem and the coupled problem with mixed LSFEM-Galerkin principle

no other forces are applied on the boundary of the structural part, it is possible to set T in Eq. (6.15)

$$T = \int_{\Gamma_{if}} -\sigma_f n \cdot \varphi \, d\Gamma_{if} \tag{6.16}$$

This then completes the coupling and the fluid-structure interaction problem is fully specified.

To clarify the numerical procedure the resulting matrix structure of the strongly coupled problem is shown in Fig. 6.1. Although the resulting tangential stiffness matrices of the subproblems are fully symmetric, the coupling destroys this symmetry.

6.5 LSFEM-LSFEM Coupling

The articles [76], [75] and [54] propose a strongly coupled least squares formulation for FSI problems. In these articles, the coupling is introduced as an additional least squares term, which seems to be the most consistent way. Hence this approach should be used in this work as well. Let $\mathcal{J}_f$ and $\mathcal{J}_s$ denote the functionals, which stem from the fluid and the structural domain. Then the coupling conditions could be added as an additional term. Thus

$$\mathcal{J}_{tot} = \mathcal{J}_f + \mathcal{J}_s + \mathcal{J}_{if} \tag{6.17}$$

with $\mathcal{J}_{if}$ being the functional measuring the residual of the coupling conditions Eq. (2.49) and Eq. (6.3). In subsection 6.3.2 the spaces in which the residual of these coupling conditions lies, were already discussed. The norms within the least squares functional should fit the spaces. As usual the different parts in Eq. (6.17) could be weighted differently. These weights will be included into the weights $\alpha_i^F, i = 0 \ldots 2$ and $\alpha_i^S, i = 0 \ldots 2$ which were already introduced for the functionals related to the subproblems.

The functional $\mathcal{J}_{if}$ is split further into a functional $\mathcal{J}_{co}$ representing the compatibility condition and a functional $\mathcal{J}_{tr}$ which implements the traction coupling condition. Then:

$$\mathcal{J}_{if} = \alpha_1^C \mathcal{J}_{co} + \alpha_2^C \mathcal{J}_{tr} \tag{6.18}$$

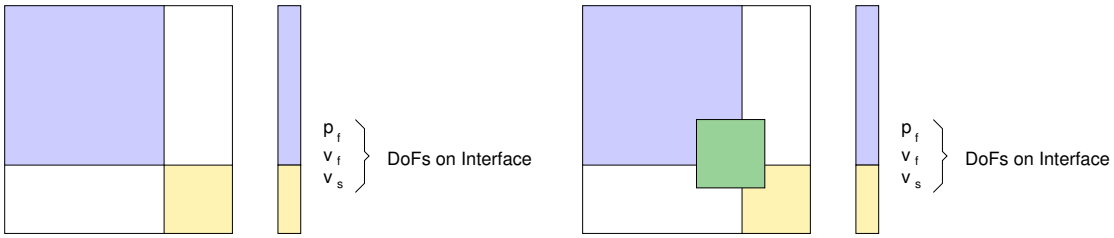

Figure 6.2. Schematic matrix structure for the uncoupled problem and the coupled problem with a unified LSFEM principle

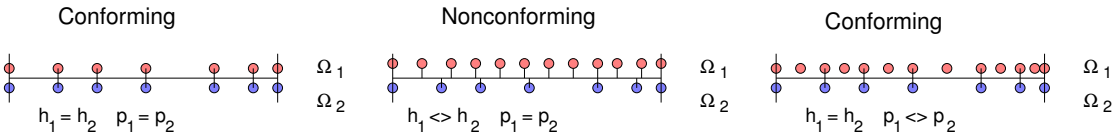

Figure 6.3. Examples of conforming and nonconforming interfaces.

where α_1^C and α_2^C are weights which will be discussed later.

Including the coupling conditions into the least squares functional keeps the symmetric structure of the resulting matrix in contrast to the previously proposed LSFEM-Galerkin coupling scheme. The effect of the coupling terms onto the resulting algebraic matrix structure is shown in Fig. 6.2.

6.5.1 Compatibility Condition

Two ways will be introduced to implement the compatibility condition. One will work only for matching interfaces between the velocity field of the fluid- and the structural domain. Here matching interfaces means that the edges of the finite elements on the interface share the same nodes at their end (cf. Fig. 6.3). If the elements on both sides have the same polynomial degree, Eq. (2.49) can be satisfied in a pointwise sense. When the polynomial degrees of the elements on the boundary differ, restricting the nodal values on the side with the higher polynomial degree leads to a pointwise fulfilment of Eq. (2.49) as well. Clearly the full convergence order will be lost in that case. The constraints which are necessary for this approach can be implemented in a way which preserves the symmetry of the resulting system of equations (cf. [4]).

As this approach will exactly satisfy the compatibility condition it provides a way to obtain reference solutions against which the solutions obtained by enforcing the compatibility condition in the least squares sense can be compared. Furthermore it offers a good way to evaluate the least squares implementation of the traction coupling condition without the influence of errors in the compatibility condition.

The other approach is the weak implementation of the compatibility condition. Referring to the subsection 6.3.2, the appropriate norm seems to be $H^{1/2}$, which leads to

the following functional:

$$\mathcal{J}_{co}(\boldsymbol{v}_f, \boldsymbol{v}_s) = \frac{1}{2}\|\boldsymbol{v}_f - \boldsymbol{v}_s\|^2_{1/2,\Gamma_{if}} \tag{6.19}$$

Due to the fractional index of this norm, the evaluation is difficult albeit possible (cf. [92], [117]). Nonetheless in this work the replacement of the $H^{1/2}$-norm by a weighted L^2 norm will be preferred. According to [117] and [26], $\|\cdot\|_{1/2,\Gamma}$ can be substituted by $h^{-1/2}\|\cdot\|_{0,\Gamma}$. The parameter h denotes again the characteristic element size.

6.5.2 Normal Traction Condition

In subsection 6.3.2 the appropriate spaces for the traction coupling condition were discussed. It is not fully clear, whether the normal stresses along fluid structure interface are in $H^{-1/2}(\Gamma_{if})$. Hence this will be explicitly assumed. Then the following functional is appropriate for the traction coupling condition:

$$\mathcal{J}_{tr}(\boldsymbol{v}_f, p, \sigma_s) = \|(-\mathbf{I}p + 2\nu_f\mathbf{D} - \sigma_s)\boldsymbol{n}_C\|_{-1/2,\Gamma_{if}} \tag{6.20}$$

Again a norm on a space with a fractional index has to be evaluated. A multilevel approach is suggested in [117]. In this work the norm will be replaced by a weighted L^2 norm. Clearly the $H^{-1/2}$ norm on the interface could be bounded by the stronger L^2 norm on the boundary without weights. But directly using the L^2 norm without appropriate weights could lead to an overweighting of that least squares term and thus to decreased accuracy (cf. [117]). Therefore $\|\cdot\|_{-1/2,\Gamma}$ will be replaced by $h^{1/2}\|\cdot\|_{0,\Gamma}$.

6.5.3 Temporal Coupling Conditions

Within the strongly coupled system of partial differential equations two types of equations can be differentiated. The constraint equations are the first type, while the evolution equations belong to the second type. In the fully continuous situation these constraints have to be satisfied at every time instance. Due to the time discretisation this does not hold in the discrete case.

Some test computations with the fully coupled least squares method showed temporal oscillations when the parameter θ was set to 0.5 for all equations. A similar effect was already observed for the Navier-Stokes equations (cf. subsection 4.3.2). First the

evolution equations will be summarised in compact form:

$$\frac{\boldsymbol{v}^f_{n+1} - \boldsymbol{v}^f_n}{\Delta t} = -\left(\theta_p \nabla p_{n+1} + (1-\theta_p)\nabla p_n\right) \tag{6.21}$$

$$-\left(\theta \nabla \times \omega_{n+1} + (1-\theta)\nabla \times \omega_n\right) \quad \text{in} \quad \Omega_f \tag{6.22}$$

$$\mathbb{C}^{-1/2}\frac{(\sigma_{n+1} - \sigma_n)}{\Delta t} = \theta_s \mathbb{C}^{1/2}\dot{\tilde{\mathbf{E}}}(\boldsymbol{v}_{n+1}) + (1-\theta_s)\mathbb{C}^{1/2}\dot{\tilde{\mathbf{E}}}(\boldsymbol{v}_n) \quad \text{in} \quad \Omega_s \tag{6.23}$$

$$\frac{\boldsymbol{v}^s_{n+1} - \boldsymbol{v}^s_n}{\Delta t} = \theta \nabla \cdot \sigma_{n+1} + (1-\theta)\nabla \cdot \sigma_n \quad \text{in} \quad \Omega_s \tag{6.24}$$

The constraints are:

$$\nabla \cdot \boldsymbol{v}^f_{n+1} = 0 \quad \text{in} \quad \Omega_f \tag{6.25}$$

$$\omega_{n+1} - \nabla \times \boldsymbol{v}^f_{n+1} = 0 \quad \text{in} \quad \Omega_f \tag{6.26}$$

$$\boldsymbol{v}^f_{n+1} - \boldsymbol{v}^s_{n+1} = 0 \quad \text{on} \quad \Gamma_{if} \tag{6.27}$$

$$\boldsymbol{n}\sigma^s_{n+1} - \boldsymbol{n}\sigma^f_{n+1} = 0 \quad \text{on} \quad \Gamma_{if} \tag{6.28}$$

The time derivative is approximated by a finite difference, which corresponds to the assumption that the temporal evolution of that unknown is linear within that time interval. The point within this interval where the steepness of the assumed linear function is evaluated is defined by the parameter θ. Taking the pressure as an example, this is also the point where the pressure acts as Lagrange multiplier for the incompressibility constraint. If the pressure is evaluated as the average of the timesteps n and $n+1$, the average will be quite accurate and does not show large oscillations. In contrast the pressure at n and $n+1$ will wildly oscillate as only the average between these two timesteps has to be correct. Setting $\theta_p = 1.0$ reduces these oscillations significantly.

The same holds for the temporal evolution of the stresses. Clearly Eq. (6.23) and Eq. (6.28) cannot be satisfied exactly in the discrete case. This leads to oscillations in the stress field. Setting $\theta_s = 1.0$ reduces these oscillations significantly but has also the disadvantage of introducing more numerical damping. Summarising these findings, the used time integration parameters will be $\theta_s = \theta_p = 1.0$ and $\theta = 0.5$.

6.6 Numerical Test of the Coupling

To numerically verify the accuracy of the used coupling procedures, a very simple configuration is used, which also possesses a stationary solution. The setup of the test case is shown in Fig. 6.4. It should simulate a Stokes flow through a channel with one elastic wall. The Stokes flow is chosen to avoid difficulties with the nonlinearity in the Navier-Stokes equations. Including this nonlinearity would not give new insights about the numerical behaviour of the coupling conditions. Another simplification was introduced by not deforming the grid in this example.

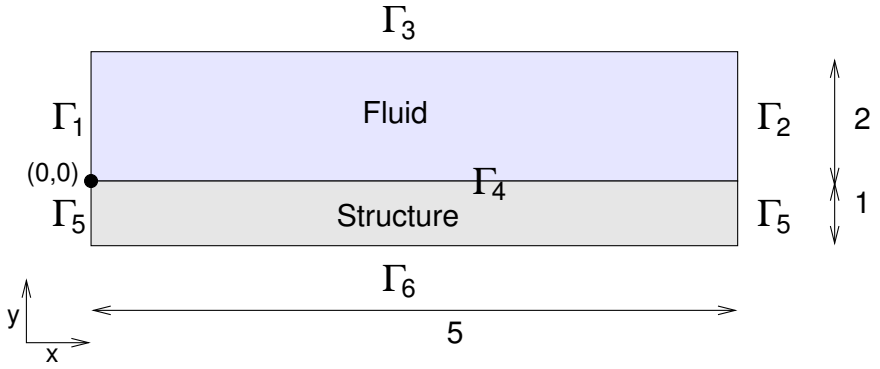

Figure 6.4. Setup of the test case for the coupling conditions

The L^2 version of the $up\omega$-formulation is used for the simulation of the fluid domain. As the $v\sigma$-formulation was the most reliable least squares formulation in the previous tests, it will be used for the structural domain.

The boundary conditions for the fluid are:

$$\boldsymbol{v}(\boldsymbol{x},t) = (2y - y^2, 0) \quad \text{on} \quad \Gamma_1 \tag{6.29}$$

$$p(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_2 \tag{6.30}$$

$$\boldsymbol{v}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_3 \tag{6.31}$$

On the boundaries of the structural domain the following conditions are imposed:

$$\boldsymbol{v}(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_5 \tag{6.32}$$

$$\boldsymbol{n} \cdot \sigma(\boldsymbol{x},t) = 0 \quad \text{on} \quad \Gamma_6 \tag{6.33}$$

Except for Eq. (6.33) all boundary conditions are implemented in the usual way. To ensure Eq. (6.33) the following least squares term is added to the functional:

$$\mathcal{J}_{bnd} = \frac{1}{2}\alpha_3^S \|\boldsymbol{n} \cdot \sigma\|^2_{0,\Gamma_6} \tag{6.34}$$

An impulsive start is used, which can be represented by the following initial conditions:

$$\boldsymbol{v}(\boldsymbol{x},0) = 0 \quad \text{in} \quad \Omega_1 \tag{6.35}$$

$$p(\boldsymbol{x},0) = 0 \quad \text{in} \quad \Omega_1 \tag{6.36}$$

$$\omega(\boldsymbol{x},0) = 0 \quad \text{in} \quad \Omega_1 \tag{6.37}$$

and

$$\boldsymbol{v}(\boldsymbol{x},0) = 0 \quad \text{in} \quad \Omega_2 \tag{6.38}$$

$$\sigma(\boldsymbol{x},0) = 0 \quad \text{in} \quad \Omega_2. \tag{6.39}$$

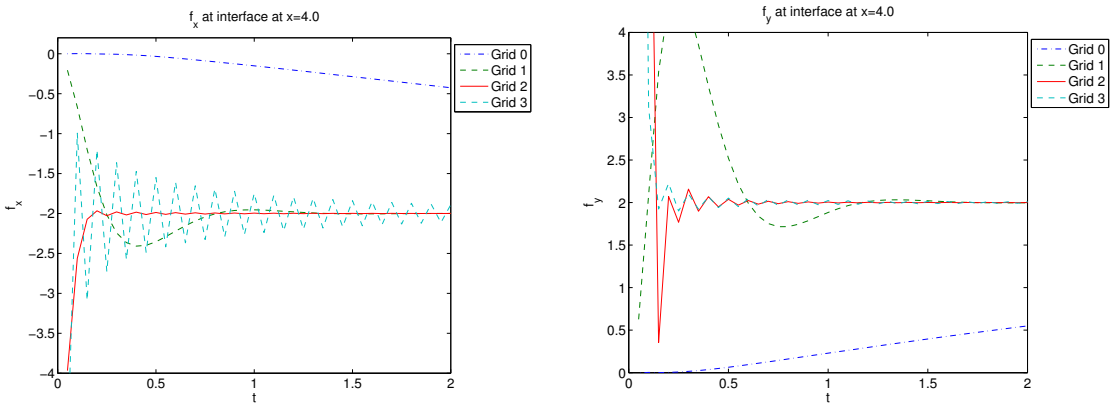

Figure 6.5. Temporal development of the forces at the point (4,0)

For the evaluation of the different possibilities to implement the coupling conditions the values of interest will be evaluated at three points on the interface $\boldsymbol{x}_1 = (1,0), \boldsymbol{x}_2 = (2,0), \boldsymbol{x}_3 = (3,0)$.

The material parameters are $E = 10^6$, $\nu_s = 0.35$ and $\rho_s = 1$ for the structure. The fluid material parameters are $\rho_f = 1$ and $\nu_f = 1$. All computations were carried out with a time step size $\Delta t = 0.05$. For the fluid the $Q3 - Q2$ element pair was selected. It was accompanied by the pair $RT2 - Q3$ on the structural side.

6.6.1 Uncoupled Solution

In this test the fluid-structure interface Γ_4 will be considered to be rigid. First the velocity unknowns along Γ_4 will be set to zero to implement a no-slip boundary condition which would lead to a completely symmetric setup. This test should fulfil two objectives. First it should give an impression of the temporal development of the forces along the interface and the dependence on the mesh refinement. Second the no-slip boundary condition will be implemented in a least squares sense to evaluate if a least squares term suffices to constrain the velocities on the no slip boundary.

Hence for the first numerical test the following boundary conditions are imposed:

$$\boldsymbol{v}_f = 0 \quad \text{on} \quad \Gamma_4 \tag{6.40}$$

$$\boldsymbol{v}_s = 0 \quad \text{on} \quad \Gamma_4 \tag{6.41}$$

and the values of the vector $\boldsymbol{f}$

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{n} \cdot \sigma_f(\boldsymbol{x}) \tag{6.42}$$

will be recorded for $\boldsymbol{x} = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_3$.

	Grid 0	Grid 1	Grid 2	Grid 3
Elements	15	60	240	960
α_0^F	1	1	1	1
α_1^F	1	4	16	64
α_2^F	1	4	16	64
α_0^S	1	1	1	1
α_1^S	1	1	1	1
α_2^S	1	1	1	1

Table 6.1. Weights for the different least squares terms in the coupled problem.

In Fig. 6.5 the evolution of the forces on the point (4,0) is shown. Actually the incompressibility constraint should immediately enforce the parabolic inflow profile throughout the whole domain. As the incompressibility constraint is only enforced weakly in the least squares sense, this does not hold for the numerical solution. But the results shown in Fig. 6.5 demonstrate that refining the spatial discretisation brings the solution closer to the exact solution.

Now the same setup will be used to evaluate the accuracy which can be obtained by weakly enforcing the velocity boundary conditions. Here the influence of the weight α_1^C and the influence of the mesh refinement will be examined. As it was already explained, the least squares functional for the Stokes equations is not H^1-coercive. Thus the weights α_1^F and α_2^F should be increased while the mesh is refined. Table 6.1 summarises the parameters of the four examined discretisations. These were obtained by uniformly refining the coarsest discretisation which consists of 15 elements.

The results for these computations are shown in Fig. 6.6. Increasing the weight of the least squares term ensuring the compatibility of the velocities clearly leads to a reduction of that functional. A more interesting observation is that the functional linearly depends on the selected weight from a certain point on.

As already pointed out, the norms of the boundary term prescribing the velocity actually require a slightly stronger norm than the used L^2 norm. Therefore the weight α_1^C has to be increased when the spatial discretisation is refined. In Fig. 6.6 this effect can be seen on the slightly chaotic behaviour of the curves for lower weights which continues up to a certain weight. This point is shifted to the regions of higher weights for the finer discretisations. Another remarkable point is the fact that the finest discretisations do not lead to the most accurate boundary approximations.

Another issue is the numerical efficiency of the least squares coupling. To analyse the effect of the weights onto the conditions number of the resulting algebraic system, a comparison between the coupling with constraints (denoted by $\alpha_1^C = \infty$) and the least squares coupling with different weights was performed. A conjugate gradient algorithm was used with Gauss-Seidel preconditioning. The iteration was stopped when the resid-

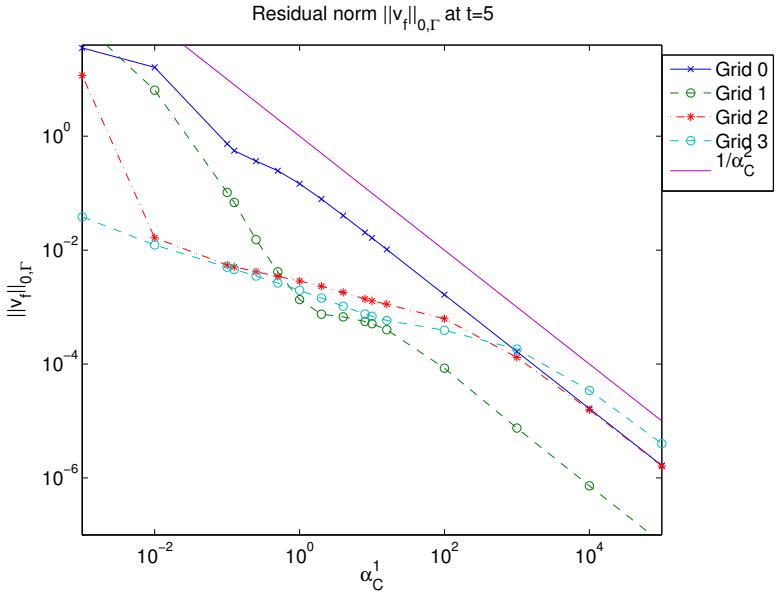

Figure 6.6. Influence of the weight α_1^C onto the residual $\|\boldsymbol{v}_f - 0\|_{0,\Gamma_4}$ for differently refined spatial discretisations.

α_1^C	1	10	10^2	10^3	10^4	10^5	10^6	∞
Grid 1	197	168	153	155	140	140	140	140
Grid 2	445	283	200	187	185	186	188	185
Grid 3	1398	597	326	260	237	239	240	239

Table 6.2. CG iterations for different grids and weights

ual was reduced by a certain factor. Table 6.2 shows the required iterations for the first time step. With increasing α_1^C the number of required iterations comes close to the one which is needed for the method using constraints. This indicates that the least squares coupling approaches the formulation using constraints when $\alpha_1^C \to \infty$.

To conclude this section, the numerical results show that it is possible to implement the velocity constraint in a weak sense. Actually this is not astonishing, as this result can be deduced from the regularity estimates for the $up\omega$-formulation, where the boundary conditions appear only in terms of norms on the trace spaces. In the discrete case the norms on these Sobolev spaces with fractional indices can be replaced by appropriately weighted L^2 norms on the boundary. If the weight α_1^C is chosen correctly, the accuracy and numerical efficiency of the least squares approach do not differ significantly from the approach using constraints. As soon as nonconforming interfaces should be treated, the least squares approach would therefore be the method of choice. In this work the inter-

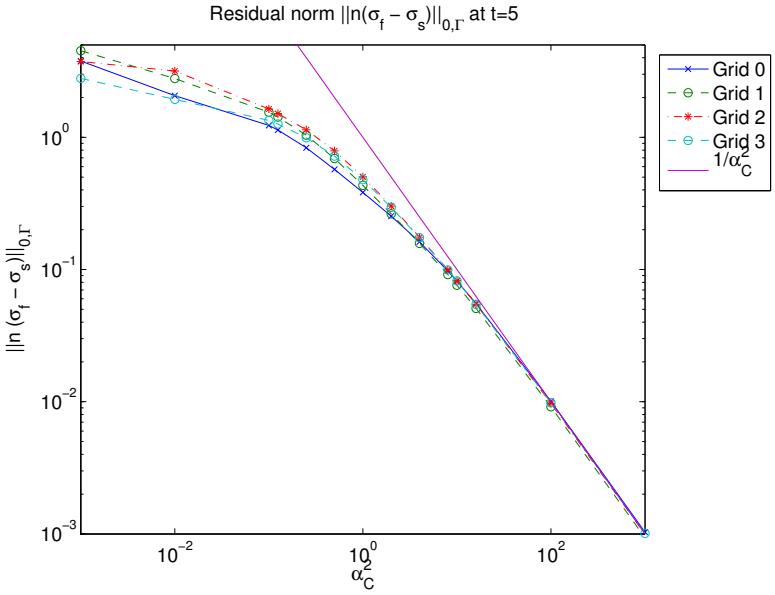

Figure 6.7. Influence of the weight α_2^C onto the residual $\|\boldsymbol{n}_C \cdot (\sigma_f - \sigma_s)\|_{0,\Gamma_4}$ for differently refined spatial discretisations.

faces will always be conforming and hence the realisation of the compatibility condition with constraints will be preferred as it avoids the determination of a proper weight α_1^C.

6.6.2 Compatibility of the Normal Tractions

Now the influence of the weight onto the normal tractions will be evaluated. For this purpose the compatibility of the velocities is implemented by the constraint mechanism to exclude any influence of inaccuracies in those coupling conditions. Again the influence of the weight, this time α_2^C, and the spatial discretisation is analysed.

The computed results are shown in Fig. 6.7. Similar to the results for the weak formulation of the compatibility condition, the curves show essentially a linear dependence between the error in the L^2 norm over the boundary and the applied weight α_2^C. This linear behaviour can be found after the weight has passed a certain value. Interestingly the curves do not differ significantly for the different spatial discretisations. A possible explanation might be the weaker $H^{-1/2}$ norm which has been replaced by the L^2 norm. It might be that this replacement influences the accuracy of the solution inside the domain due to the improper weighting of the least squares terms, while the accuracy of the coupling conditions remains unaffected.

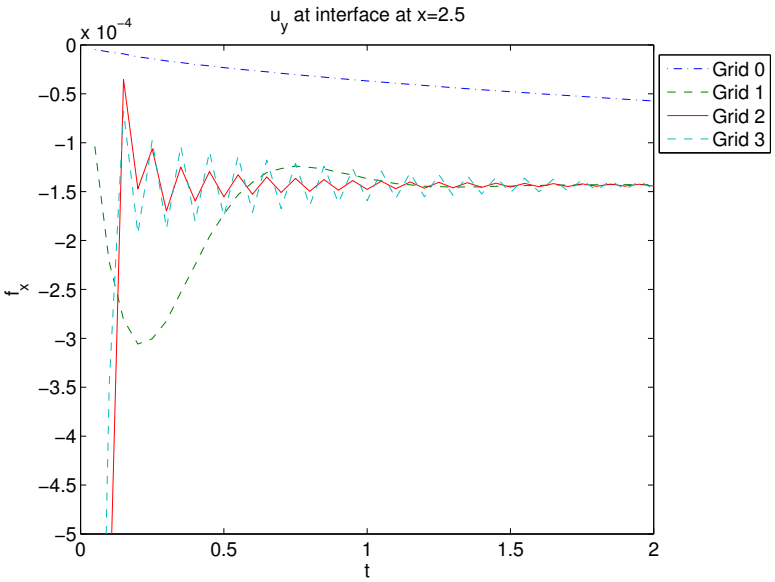

Figure 6.8. Vertical displacement u_y of the elastic structure. Full least squares formulation.

6.6.3 Full LSFEM-Coupling

In contrast to the previous test, where Eq. (2.49) was implemented by constraining the nodal values on the side with the lower polynomial degree, this time both coupling conditions will be included as least squares terms. The results of the previous test, which are shown in Fig. 6.6 and Fig. 6.7 motivated the choice of the weights $\alpha_1^C = 10000$ and $\alpha_2^C = 100$. These weights are at the point where the dependence between the residual of the coupling conditions and the selected weight becomes approximately linear. Furthermore for the compatibility condition this weight seems to be close to the point, where the number of conjugate gradient iterations becomes stable (cf. Table 6.2).

In Fig. 6.8 the vertical displacement of the centre of the elastic bar which is located below the fluid is shown. The qualitative behaviour is similar to the development of the forces on the bottom boundary (cf. Fig. 6.5) in the uncoupled problem. Both diagrams show that the mass conservation is enforced too weak on the coarser discretisations. The flow field slowly develops and so do the forces on the boundary. When the spatial discretisation becomes finer and the weights α_1^F and α_2^F are increased in compliance with the theory, the qualitative behaviour comes much closer to the expected behaviour. An unwanted side effect is that the temporal oscillations increase. But the example demonstrates that the general idea of formulating the strongly coupled problem in a single least squares functional does work.

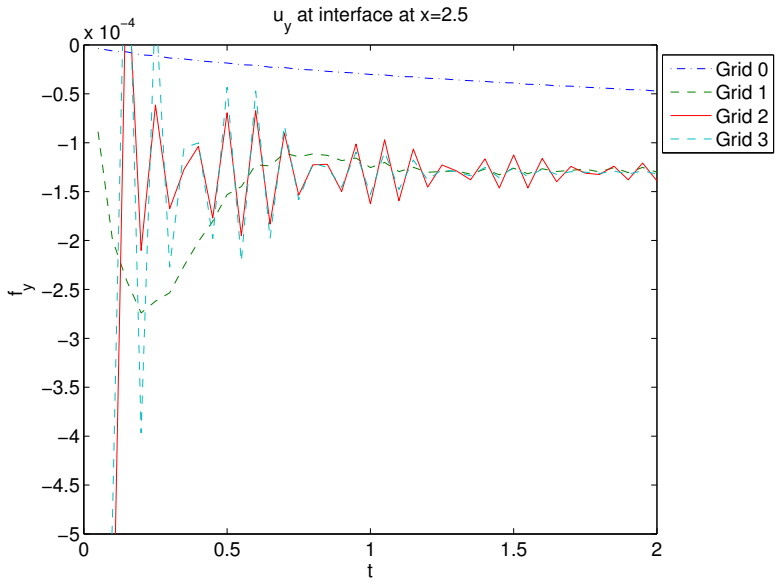

Figure 6.9. Vertical displacement u_y of the elastic structure. Galerkin formulation for the structure, LSFEM formulation for the Stokes equations.

6.6.4 LSFEM-Galerkin Coupling

The LSFEM-Galerkin Coupling was tested with the same setup. To weaken locking effects in the Galerkin formulation for the structural part, Q2 elements were used to approximate the displacement- and velocity field. For the Stokes formulation the weights which are shown in Table 6.1 were used to avoid differences within the least squares fluid part.

Again the vertical displacements of the structural part were recorded. The results are shown in Fig. 6.9 and are in good agreement with the data which was computed with the full least squares formulation in the previous paragraph. The main difference lies in the stronger temporal oscillations which are probably caused by the non existent numerical damping of the Newmark scheme for the structural part. Here the inherent dissipation of the least squares formulation clearly has a damping effect on the temporal oscillations. From a practitioners point of view the weaker damping is clearly an advantage of the LSFEM-Galerkin Coupling. Especially in the interesting cases which are on the limit of physical stability, the damping of the least squares method might be too strong.

6.7 Physical Weighting

Preliminary numerical tests for more difficult FSI problems (i.e. the example in chapter 7) showed that an improper weighting between the parts of the functional which belong to the fluid domain and those which belong to the structural domain, leads not only to instabilities but also to wrong numerical solutions. Hence in this section the correct weights with respect to the physical units will be derived.

The unit for the squared equation residual will be selected as the one of the structural part (cf. section 5.9) which was $[N/m^2 s]$. In section 4.6 unit of the residual belonging to the Navier-Stokes equations was unified to $[m/s^2]$ which after squaring becomes $[m^2/s^4]$. Hence the weights derived in section 4.6 have to be modified such that they transfer $[m^2/s^4]$ into $[kg/s^3 m]$ which is another writing of $[N/m^2 s]$. This can easily be achieved by multiplying the weights $\alpha_0^F \dots \alpha_2^F$ with $\rho_f \Delta t$.

For the boundary residuals the integration is carried out over a lower dimensional domain. Therefore the squared residual of the coupling conditions finally has to have the unit $[N/ms]$ or $[kg/s^3]$. As it was already explained in the previous parts, the residual of the compatibility condition should be evaluated in the $H^{1/2}$ norm. To replace this norm by a weighted L^2 norm in the finite dimensional case, the L^2 norm should be weighted by $h^{-1/2}$ (cf. [117], [26]). The squared residual of Eq. (2.49) has a unit of $[m^2/s^2]$. Using

$$\alpha_1^C = \rho_s \cdot \frac{\bar{L}\bar{\boldsymbol{v}}}{h} \tag{6.43}$$

as weight, the unit of the weighted squared equation residual of Eq. (2.49) becomes:

$$\rho_s \cdot \frac{\bar{L}\bar{\boldsymbol{v}}}{h} \left[\frac{kg}{sm^2}\right] \cdot \left[\frac{m^2}{s^2}\right] = \left[\frac{kg}{s^3}\right] \tag{6.44}$$

which after integration over the interface has the conforming physical unit $[kgm/s^3]$.

The squared residual of Eq. (2.48) has a unit of $[N^2/m^4]$ or $[kg^2/s^4m^2]$. This time the factor which must be introduced to replace the weaker $H^{-1/2}$ norm by the L^2 norm in the finite dimensional case is $h^{1/2}$. Therefore a suitable weight would be:

$$\alpha_2^C = \frac{h}{\rho_f \bar{L}\bar{\boldsymbol{v}}}. \tag{6.45}$$

Combining this with unit of the squared equation residual yields:

$$\frac{h}{\rho_f \bar{L}\bar{\boldsymbol{v}}} \left[\frac{m^2 s}{kg}\right] \cdot \left[\frac{kg^2}{s^4 m^2}\right] = \left[\frac{kg}{s^3}\right] \tag{6.46}$$

which after integration over Γ_{if} is compatible with the units of the other equation residuals.

An open question is which specific mass density has to be used in Eq. (6.44) and Eq. (6.46) respectively. Basically four choices are possible. The first one was inserted into the equations above. Clearly the specific mass densities could be exchanged, which gives the second possibility. Another option is to chose one of the densities for both equations. Finally an average of both densities can be taken.

Interestingly the factor $\bar{L}\bar{\boldsymbol{v}}/h$ appears in almost all weights with different powers. Hence it will be denoted by α_M where the M should stand for *mesh* because it seems to be related to the used discretisation. Summarising these findings, the following weights will ensure a physically meaningful and dimension independent formulation for the $up\omega$-formulation coupled with the $v\sigma$-formulation:

$$\alpha_M = \frac{\bar{L}\bar{\boldsymbol{v}}}{h} \tag{6.47}$$

$$\alpha_0^S = 1 \tag{6.48}$$

$$\alpha_1^S = 1 \tag{6.49}$$

$$\alpha_2^S = 0 \quad \text{or} \quad \frac{\Delta t}{\rho_s \bar{L}^2} \tag{6.50}$$

$$\alpha_0^F = \rho_f \cdot \Delta t \tag{6.51}$$

$$\alpha_1^F = \rho_f \cdot \Delta t \cdot (\alpha_M)^2 \tag{6.52}$$

$$\alpha_2^F = \rho_f \cdot \Delta t \cdot (\alpha_M)^2 \tag{6.53}$$

$$\alpha_1^C = \rho_s \cdot \alpha_M \tag{6.54}$$

$$\alpha_2^C = (\rho_f \cdot \alpha_M)^{-1} \tag{6.55}$$

6.8 Computation of Grid Deformation and Grid Velocity

The grid deformation can be either included into the strongly coupled algorithm, what would then really lead to a fully coupled algorithm, or be treated separately. In a series of publications, which also consider the application of the least squares FEM for the solution of FSI problems, this fully coupled approach is proposed (cf. [54], [76]). Nonetheless, most algorithms for FSI problems nowadays do not include the mesh deformation into the problem, as this could cause problems (cf. [127]). Normally, if the mesh deformation is updated after each nonlinear iteration, this does not lead to problems. Therefore we also use this approach, and deform the computational grid according to the equations of linear elasticity with zero inertia after each nonlinear iteration (cf. [127],[6]).

Usually the mesh is refined around the embedded structure to resolve the boundary layer more accurate. Unfortunately, the largest mesh deformations usually appear also in the direct neighbourhood of the structure, where the elements are comparatively small. If the same elastic modulus is assigned to all elements within the domain, larger

deformations could lead to an improper deformation of the smaller elements (cf. [127]). Improper deformation in this context means that the Jacobian of the mapping becomes negative which inevitably leads to a breakdown of the numerical procedure.

To overcome this difficulty, a larger elastic modulus is assigned to smaller elements. A similar procedure was proposed in [127]. Here the following formula is used to determine the elastic modulus:

$$E = \frac{1}{h^2} \tag{6.56}$$

where h is the diameter of a circle circumscribing the whole element.

For the stability of the coupled procedure it is important that the mesh velocities, which go into the ALE-formulation, are in compliance with the structural velocities. That means the accuracy of the time discretisation procedure for the mesh velocities should match the accuracy of the time discretisation of the structural velocities.

Therefore the mesh velocities are computed with the Newmark procedure for the LSFEM-Galerkin coupling. This ensures that the velocities along the fluid-structure interface are the same in the temporal direction. Once the mesh velocities have been computed by solving a system of equations, the mesh deformations can easily obtained.

It should be noted that the LSFEM-Galerkin formulation still is not fully conservative due to the nonmatching grid deformation which is implemented by bilinear finite elements. A simple solution of this problem cannot be found due to two interdependent reasons. To ensure that the grid deformations match the velocities, the grid deformation must be computed by higher order elements. These higher order grid deformations would then require a higher order time discretisation for the fluid domain to satisfy the GCL. Both problems could theoretically be solved but are beyond the scope of this work.

For the LSFEM/LSFEM coupling the problem becomes more difficult. Due to the absence of the structural deformations in the least squares formulation, the displacements have to be recovered. Either the displacements can be recovered from the stress field (cf. section 5.8.5) at each time instance or they can be found be integrating the velocity field over time. Both approaches have been tested for the example in chapter 7.

The advantages of using the velocities to obtain the displacement are the efficiency and the consistency because the grid velocities on the interface really match the true velocities. But it came out that the displacements which are computed with this approach have a tendency to drift away. This is probably another consequence of the lack of full conservation and leads to the destruction of the computational grid after some time.

Clearly the recovery of the displacements from the stress field requires the solution of an additional system of equations in each time step or nonlinear iteration. Furthermore, it is not guaranteed that the grid deformation on the interface is compatible with the velocities on the interface. Despite these theoretical disadvantages, that approach proved to be computationally more stable. After the grid deformations $\boldsymbol{u}_{n+1}$ have been computed, the grid velocities $\boldsymbol{v}_{n+1}$ are derived from the grid deformations by a "reverse"

trapezoidal rule.

$$\frac{\boldsymbol{u}_{n+1}-\boldsymbol{u}_n}{\Delta t}=\frac{1}{2}\boldsymbol{v}_n+\frac{1}{2}\boldsymbol{v}_{n+1}\Rightarrow 2\frac{\boldsymbol{u}_{n+1}-\boldsymbol{u}_n}{\Delta t}-\boldsymbol{v}_n=\boldsymbol{v}_{n+1} \tag{6.57}$$

6.9 Summary

This chapter was devoted to the possible ways to include the coupling conditions into a numerical scheme which should find the solution of the strongly coupled FSI problem. Currently most of the literature about this problem considers the Galerkin variational scheme. Therefore, the intention of this chapter was to support the validity of the used coupling schemes with some computational test.

First a short review of an article about the existence and uniqueness of solutions for the Stokes equations which are coupled with the equations of linear elastodynamics was given. It was intended to support the following section which discussed the appropriate functional setting for a FSI problem. Actually a mathematically precise analysis would have required at least some stability estimates. These are currently not known for the least squares formulation and are excluded from this work.

Nonetheless two approaches for a strongly coupled method were proposed. The first one could be used to couple the Galerkin formulation of the equations of linear elastodynamics with the $up\omega$-formulation for the Stokes and Navier-Stokes equations. It is essentially a combination of the natural boundary conditions for the traction coupling condition and a weak formulation of the compatibility condition. Its disadvantage is the loss of symmetry of the resulting algebraic system of equations. Attempts to couple the Galerkin formulation with the least squares formulation by least squares terms were not successful.

Therefore, the second proposed method is focused on coupling two least squares methods for the subproblems. This approach is currently not mathematically verified but has been used by other authors with success (cf. [54], [76], [75]). It recasts the full coupled system into an equivalent minimisation problem which then is solved by the least squares method.

Finally, the largest part of this chapter considered a simple model FSI problem which consisted of a channel flow with one elastic wall. With this example the coupling conditions were examined separately. After that the fully coupled model problem was solved once with the pure least squares approach and once with the proposed LSFEM-Galerkin coupling. The results of both computations were in good agreement and confirmed that the coupling schemes at least produce the same results which look physically reasonable.

After that a short informal discussion of the other components which are necessary to solve real FSI problems was presented. A special difficulty with the least squares formulations for the structural part is the absence of the displacements which makes the recovery of these fields necessary.

To conlude this section, two new coupling schemes were proposed to couple least squares formulations with other least squares formulations and with Galerkin formulations. It came out that a pure least squares formulation is able to get accurate results at least for the used model problem.

7 Example

Besides the computations which were already presented in the chapters for the subproblems, here a fully coupled computation for a larger model problem will be presented. This model problem has been computed by several other authors (cf. [128], [79], [127], [119]) with different methods which allows a comparison of the results which were obtained with the LSFEM. After a short description of the problem setup, results for the LSFEM-Galerkin coupling scheme are presented. Finally, results which were computed with the pure least squares formulation will be shown.

7.1 Problem Setup

This FSI problem was proposed by Wall in [128] and since then has been solved by some other authors (cf. [97], [127]). As these authors obtained very similar results, it is reasonable to use their results as a reference.

Geometrically the setup of this example consists of an elastic flap, which is attached to a fixed block. The block and the flap are then embedded into a channel flow with uniform inflow velocity. This setup is shown in Fig. 7.1. Due to an imposed disturbation or numerical errors, the initial symmetric flow pattern will become non symmetric. At the downstream edges of the fixed block, vortices will develop and induce an oscillation of the elastic plate. As a very detailed analysis in [127] showed, the vortex shedding frequency and the eigenfrequencies of the plate determine wheter the plate displacement shows chaotic behaviour in time or if a regularly oscillating solution occurs.

In [127] two different material parameters were used for the plate and the fluid. For the fluid the material parameters were the same in both cases. The density was selected to be $\rho_f = 1.18 \cdot 10^{-3} g/cm^3$ and the shear viscosity $\eta_f = 1.82 \cdot 10^{-4} gcm/s$. To obtain two different Reynolds numbers the inflow velocity was set to $51.3 cm/s$ in the first case and to $31.5 cm/s$ in the second case.

For the structural part two different sets of material parameters were selected in [127]. In the first case (plate I), the structure has a density of $\rho_s = 0.1 g/cm^3$, an elastic modulus of $E = 2.5 \cdot 10^6 gcm/s^2$ and a poisson ratio of $\nu_s = 0.35$. The second set of material parameters was $E_s = 2.0 \cdot 10^6 gcm/s^2$, $\rho_s = 2.0 g/cm^3$ and $\nu_s = 0.35$ (plate II). The plane stress assumption was used in both cases.

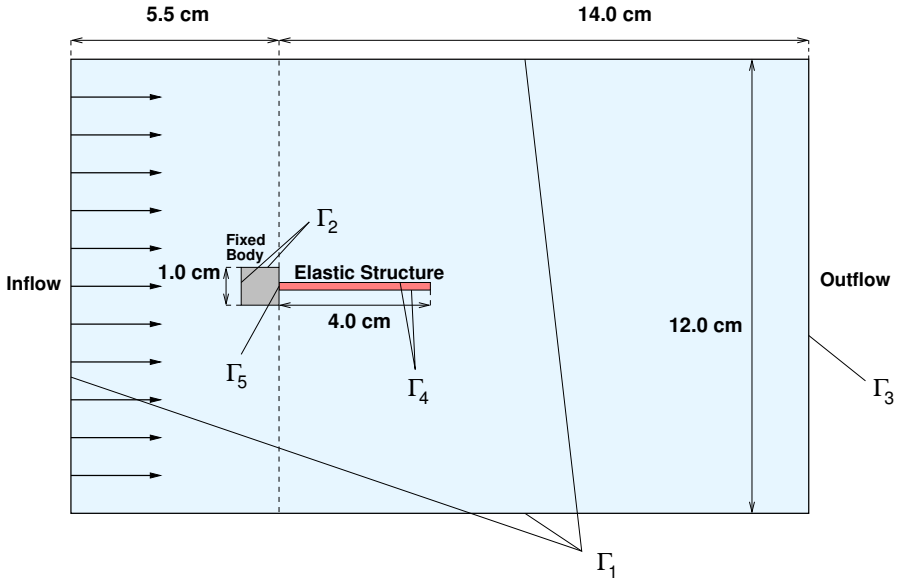

Figure 7.1. Schematic setup of the FSI problem.

As the plane stress assumption was not implemented in the least squares formulations, the LSFEM-Galerkin coupled formulation will be used to compute solutions with the plane strain assumption. This solution can then be compared with the solution obtained by the pure LSFEM formulation. Actually the differences between the plane strain and the plane stress assumption are comparatively small in this case. The eigenfrequencies were computed for both assumptions with ANSYS (cf. [2]) using the PLANE83 element with enhanced strain. With the plane stress assumption the computed first three eigenfrequencies for plate I were $3.031Hz, 18.982Hz, 53.098Hz$. The frequencies were $3.24Hz, 20.286Hz, 56.738Hz$ when the plane strain assumption was used.

As the purpose of this section is primarily the proof of concept, only the problem with plate I and an inflow velocity of $51.3cm/s$ was computed. The boundary conditions for the fluid domain were:

$$\boldsymbol{v}_f = (51.3, 0)^T \quad \text{on} \quad \Gamma_1 \tag{7.1}$$

$$\boldsymbol{v}_f = 0 \quad \text{on} \quad \Gamma_2 \tag{7.2}$$

$$p = 0 \quad \text{on} \quad \Gamma_3 \tag{7.3}$$

On the boundaries of the structure the following conditions were prescribed:

$$\boldsymbol{v}_s = 0 \quad \text{on} \quad \Gamma_5 \tag{7.4}$$

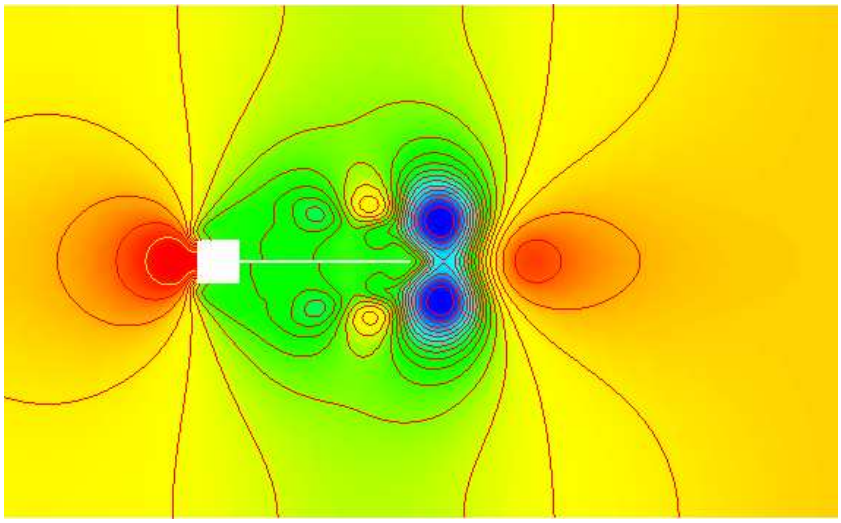

Figure 7.2. Pressure field $t = 0.5$

Along Γ_4 the coupling conditions were implemented by the procedures described in chapter 6. For the Galerkin formulation also the displacements were fixed on Γ_5:

$$\boldsymbol{u}_s = 0 \quad \text{on} \quad \Gamma_5 \tag{7.5}$$

While most of the above mentioned authors triggered the break of symmmetry by increasing the inflow velocity on single nodes on Γ_1 in the first time step, here a slightly different approach was used. A small downward acceleration $a_y = -0.1cm/s^2$ was applied as a body force onto the elastic flap. This force deforms the solid and hence introduces a small asymmetry into the fluid domain.

7.2 LSFEM-Galerkin

This section will show results which were obtained by coupling a Galerkin formulation for the structure with the $up\omega$-formulation for the fluid domain. The weights for the $up\omega$-formulation were the same as those which were used for the pure LSFEM formulation and can be found in Table 7.1. The time step size was $\Delta t = 0.01$.

In a series of tests the different element pairs $Q4 - Q3$, $Q3 - Q2$ and $Q2 - Q1$ for the LSFEM fluid part were coupled with a Galerkin formulation for the structural part using a 25 node $Q4$ element. This high order element was used to exclude inaccuracies in the structural part. Due to the low number of elements in the structural domain, the additional numerical effort for this high order element is negligible.

The pressure field at three time instances which is shown in Figs. 7.2–7.4 shows qualitatively the expected results. On the front corners of the fixed block some vortices

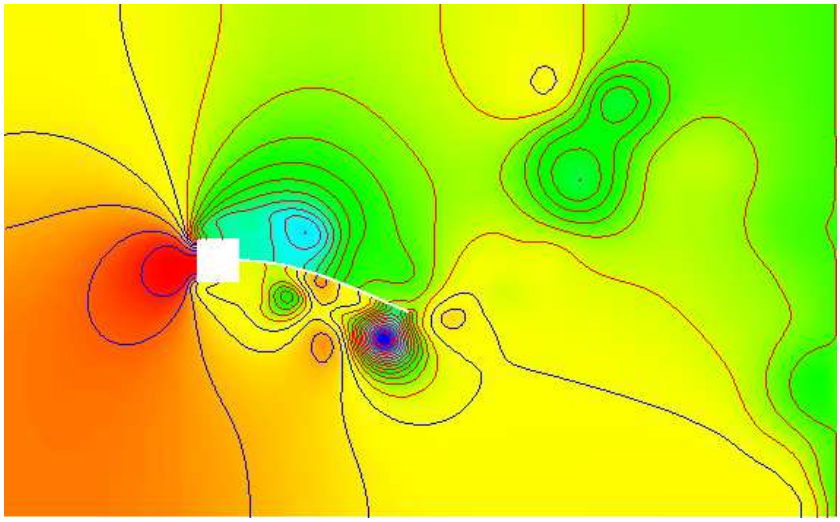

Figure 7.3. Pressure field $t = 10.12$

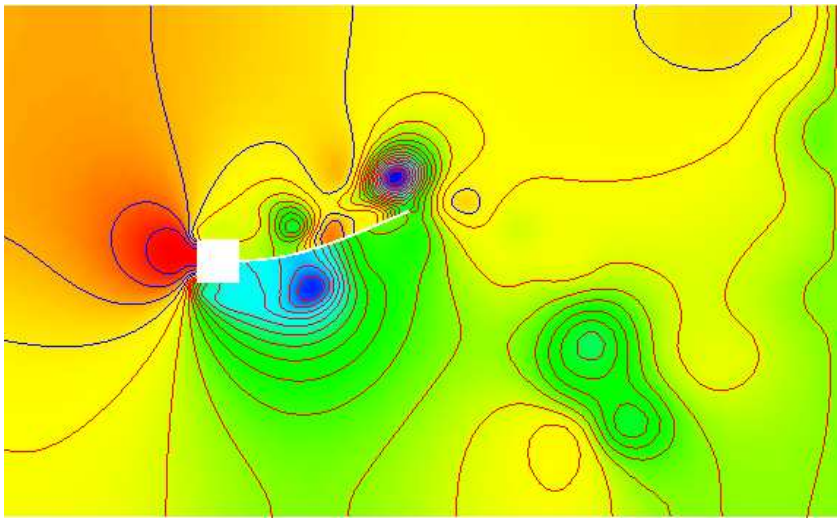

Figure 7.4. Pressure field $t = 10.28$

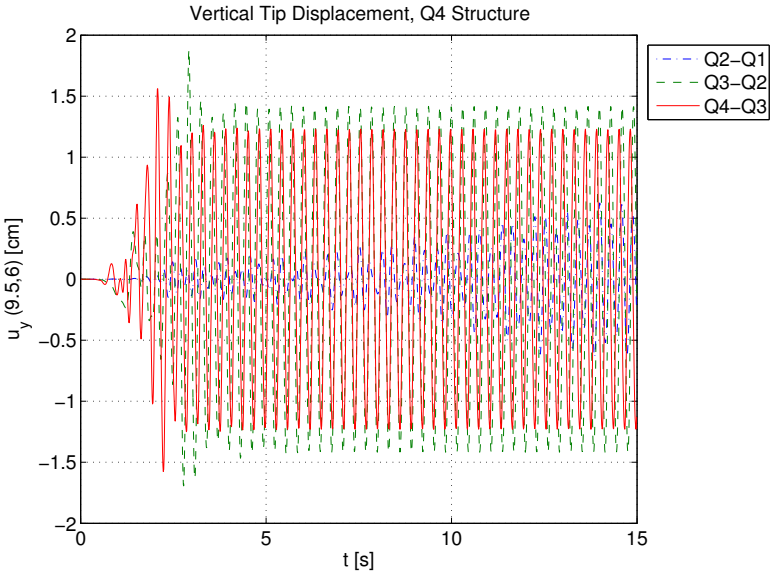

Figure 7.5. Solutions computed with the different LSFEM elements, Plane Strain

develop and are transported to the back of the elastic bar (cf. Fig. 7.2). After some time the solution reaches the quasistationary state where the flap oscillates with its first eigenfrequency (cf. Fig. 7.3 and Fig. 7.4).

Fig. 7.5 shows results which were obtained with the LSFEM-Galerkin formulation and different element pairs. Comparing these with results from [79] which are shown in Fig. 7.6 shows a good correspondence for the $Q3-Q2$ and the $Q4-Q3$ element. Interestingly the amplitudes in the quasistationary state are slightly lower for the $Q4-Q3$ element than those which were found with the $Q3-Q2$ element. The reasons for the effect are not completely clear.

Looking at the spectrum, which is shown in Fig. 7.7, shows a peak at the first eigenfrequency of the structure. According to detailed study of this test case in [127], this is exactly the expected behaviour. The structural part begins to influence the fluid part by its motion, and hence the eigenfrequency of the coupled system is essentially dominated by the first eigenfrequency of the structure. An interesting detail is the fact that the results for computation with the $Q2-Q1$ element for the structure also show a peak in the spectrum at the correct frequency. Furthermore the time series show a steadily increasing amplitude of the oscillations. Thus on the long term the $Q2-Q1$ element might reach the quasistationary state as well. Nonetheless the use of the higher order elements seems to be preferable.

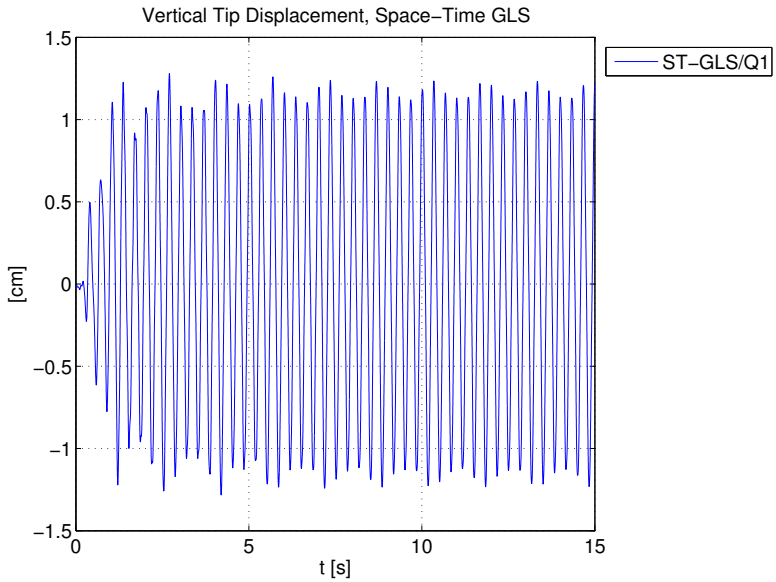

Figure 7.6. Solution with the space-time GLS (from [79])

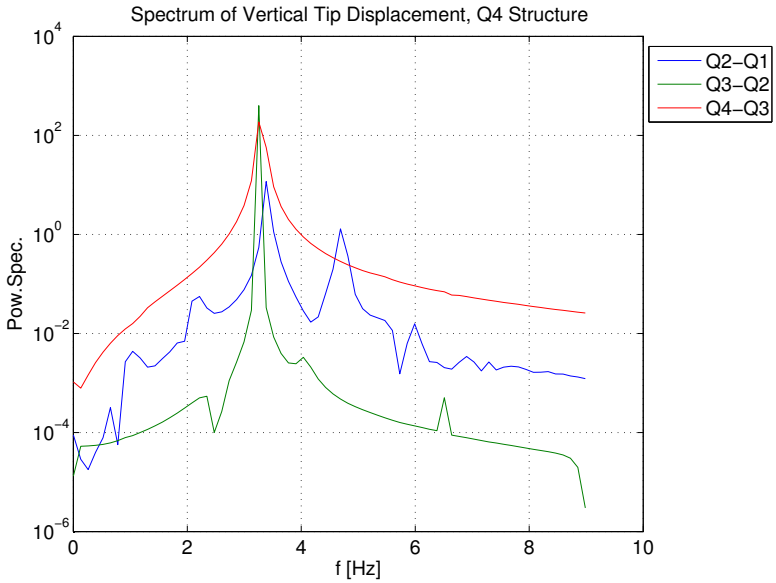

Figure 7.7. Power spectrum of vertical tip displacement for different fluid elements, LSFEM-Galerkin

Weight	Value
α_1^C	20.4
α_2^C	4.15
α_0^S	1
α_1^S	1
α_2^S	0.1
α_0^F	$1.18 \cdot 10^{-5}$
α_1^F	0.49
α_2^F	0.49

Table 7.1. Weights for the strongly coupled least squares formulation

7.3 LSFEM/LSFEM

The weights for the LSFEM/LSFEM coupled computations are computed by Eqs. (6.47–6.55). For the material parameters of the example this leads to the weights which are shown in Table 7.1. The length of one edge of the block was used as characteristic length $\bar{L} = 1$. For the characteristic element length $h = 0.25$ was assumed. The charcteristic velocity $\bar{v}$ was set to the inflow velocity of 51.3. Again a timestep size of $\Delta t = 0.01$ was used for the computations.

Three different element pairs, $RT1 - Q1$, $RT2 - Q2$ and $RT3 - Q3$ were tested. For the fluid side only the most accurate element pair $Q4 - Q3$ was used. While the traction coupling condition was imposed in the least squares sense, the compatibility condition was satisfied in the pointwise sense by restricting the nodes of the higher order side (due to the choice of elements this was always the fluid side). Besides this difference also the weight for the symmetry constraint was slightly increased to $\alpha_2^S = 1.0$ in contrast to the values shown in Table 7.1.

The temporal development of the vertical displacement at the tip of the elastic structure is shown in Figs. 7.8–7.9. Major differences in the structural response can be seen between the three computations. In the computation with the lowest order structural element, a much higher frequency of the oscillation can already be seen in the time series data. The amplitude is one order of magnitude smaller than the expected amplitude. This leads to the conclusion that the stronger numerical damping of the low order formulation avoids larger deformations of the structure which could lead to an interaction with the fluid. Hence the vortex shedding frequency of the fluid part which is about 6.2 Hz according to [127] dominates the structural response. Both higher order elements come closer to the expected behaviour. Here the resonance frequency of the structure is activated and starts to dominate the overall phenomena due to the large displacements. While the amplitudes are still too small for the $RT2 - Q2$ element, the structural response with the $RT3 - Q3$ element shows additional frequencies.

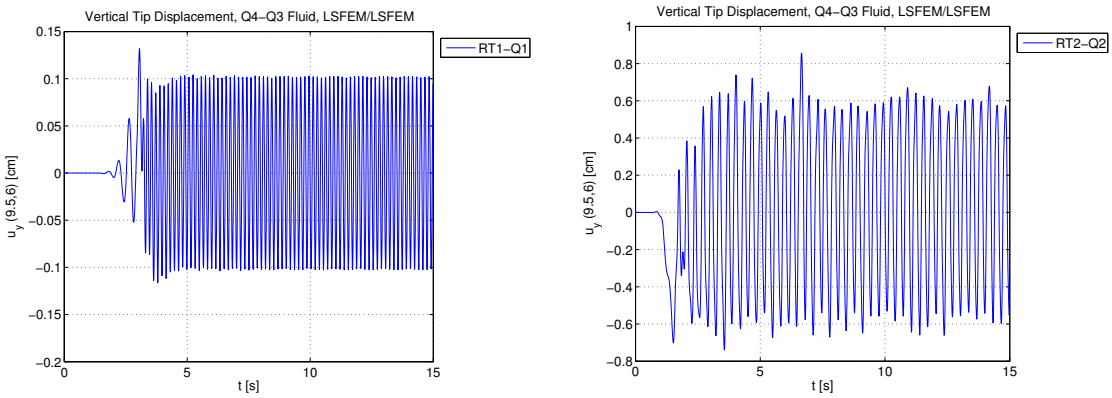

Figure 7.8. Vertical tip displacement for the $RT1 - Q1$ element and the $RT2 - Q2$ element.

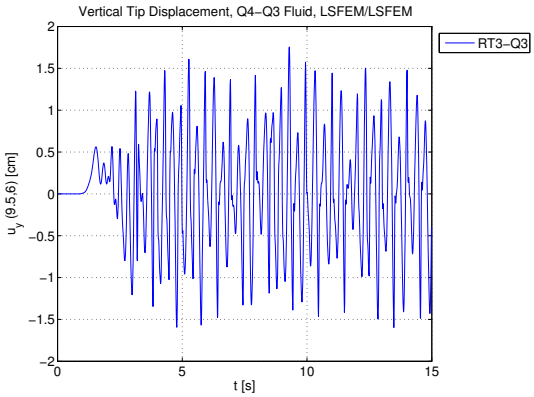

Figure 7.9. Vertical tip displacement for the $RT3 - Q3$ element.

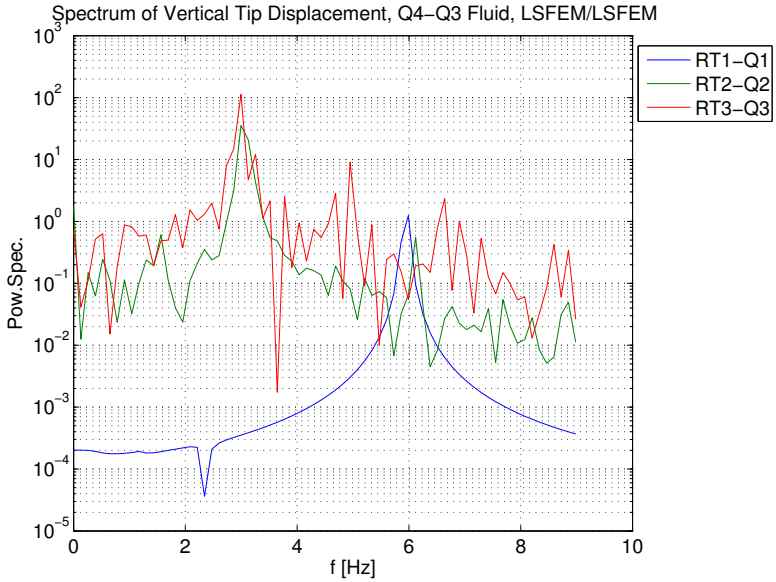

Figure 7.10. Power spectrum of vertical tip displacement, for different structural elements, LSFEM/LSFEM

To get more information about the structural response, the spectrum of the vertical tip displacement was computed by a FFT with length 768, starting from timestep 500. The resulting power spectrum is shown in Fig. 7.10. It confirms the findings of the time series interpretation. In the $RT1 - Q1$ computation, the structure simply follows the vortex shedding frequency. The spectrum has a peak at 3 Hz for both higher order structural elements. Interestingly the $RT3 - Q3$ element also shows a peak at about 5 Hz which explains the slightly chaotic impression of the time series. While the 3 Hz peak is in good agreement with the results for the LSFEM-Galerkin formulation and the results from literature, the reason for the 5 Hz peak remains unclear. An analysis of the computed structural deformations showed that structure exhibits more complex deformation patterns. Wheter these can be attributed to a lower bending stiffness or the missing geometric nonlinearities is not clear yet. It is also possible that the missing guaranteed conservation along the fluid structure interface is responsible for the more chaotic structural response of the pure least squares formulation.

7.4 Summary

In this chapter the LSFEM-Galerkin scheme and the pure least squares formulation were tested with a more complex example which is known to show complex transient behaviour in dependence of the material parameters and the boundary conditions. When

high order elements are used for the fluid part, the LSFEM-Galerkin method is able to find the solution which was reported by other authors. This fact leads to the conclusion, that the LSFEM-Galerkin formulation works quite reliable.

The results which were obtained with the pure least squares formulation differed from the results which were reported in literature. For the lower order structural elements, this difference could be explained by the high numerical damping of these elements. The reasons for the slightly more chaotic behaviour which was seen when the higher order structural elements were used, remain unclear. Nonetheless the spectrum of the solution for the two higher order structural elements was in good agreement with the expected solution. This indicates that a pure least squares solution procedure is generally possible. Further examinations of this problem are clearly required.

8 Conclusion

This chapter will summarise the findings of the previous chapters and provide a short outlook on possible future directions for the least squares FEM applied to FSI problems.

8.1 Summary

After a brief introduction of the models used to describe FSI problems, the theoretical foundation of the LSFEM was summarised. The theory was illustrated with some examples from literature which showed some of the usual techniques. Additionally a discussion of the relevant literature about each of the addressed problems was provided with remarks about gaps in the available literature. Especially the treatment of transient problems with the LSFEM currently has a thin theoretical foundation. The FSI problems which were considered in this work involve three subproblems, the fluid, the structure and the interface. Each of this problems was examined with respect to the applicability of the LSFEM in a single chapter.

First the application of the LSFEM for the fluid part was examined. In this area already a vast amount of literature is available. Essentially three main formulations in different variations have been proposed so far. As the review of the available literature revealed an analysis of formulations for the Navier-Stokes equations for transient flows is not available yet. Such analysis would have been a work on its own. Hence the approach in this work was to examine the properties of the proposed formulations computationally.

An important point which has not yet been fully clarified by literature is the quality of mass conservation in the least squares methods for the Stokes and Navier-Stokes equations. Therefore two setups were used to obtain new results regarding this problem. In the presence of sharp corners, the LSFEM can produce completely wrong results as the corner singularities induce very high squared norms of the divergence term which have a negative effect on the solution.

To get some computational convergence results for the least squares formulation for the Navier-Stokes equations for transient flows, numerical solutions of Taylor vortex were computed. It was possible to computationally confirm the second order accuracy in time, which was stated by many authors.

The last sections of the fluid chapter covered numerous computations which were performed with the different least squares formulations for a popular benchmark problem. These computations close a gap in the available literature as many articles only consider a single formulation and emphasise the benefits of that specific formulation. A consistent comparison of the different formulations on a single problem under the same circumstances was to our knowledge not done yet. The main result is that the negative norm methods perform very well, while the weighted L^2 formulations seem to suffer from their lack of full H^1 coercivity. On nonuniform grids this makes the determination of proper weights quite challenging. But for higher order elements also the weighted L^2 formulations produced highly accurate solutions for the stationary as well as for the transient problems. A drawback of the negative norm methods is the lack of literature about suitable preconditioners which work well with significant convection.

Originally the assumption was that the structural part could be handled without difficulties by the LSFEM. Several publications consider the least squares FEM for the problem of linear elasticity. But to our knowledge no publication exists about the LSFEM for linear elastodynamics. Nonetheless the extension of the formulations for elastostatics seems straightforward. Unfortunately, the numerical tests with these intuitively developed formulations were mostly unsuccessful.

This motivated the development of a simple theory for the LSFEM applied to transient problems. Starting with the wave equation which is of the same type as the equations of linear elastodynamics, the stability of several first order formulations was examined. It came out that the stability of the transient formulations is directly connected to the coercivity and continuity of the bilinear form which is related to one timestep. A new theorem was developed which then shows the stability of the solution for an arbitrary number of timesteps.

With these results it was possible to show the stability of two new least squares formulations which could be used for real problems (a third formulation was shown to be stable but only with impractical boundary conditions). The first formulation is the adaptation of a formulation for linear elasticity which has become very popular in the last years. It uses the full displacement gradient as problem unknown and recovers the displacements in a second step. By introducing the velocities as additional unknowns it was possible to derive a formulation for linear elastodynamics. For this formulation it was then possible to show full H^1 coercivity with the standard elliptic theory. Another approach was used for the second formulation. It is also based on a formulation for linear elastostatics. Formulating the constitutive law in terms of the stress and strain rates and introducing the velocity instead of the displacements gave a formulation for elastodynamics. Due to the similar structure of the equations it was possible to modify the coercivity proof for the elastostatics formulation to show coercivity of the new formulation. In contrast to formulation using the displacement gradient, coercivity can only be shown for $H(\mathrm{div};\Omega)^d \times H^1(\Omega)^d$.

Both formulations were tested with several examples to verify the stability and to test how well the conservation of energy is satisfied. With the standard LSFEM full conservation cannot be expected. The numerical tests show that for both formulations the energy conservation depends on the temporal and spatial discretisation which actually had to be expected as the LSFEM will be conservative in the limit $\Delta t, h \to 0$.

A thorough error analysis of both formulations was not carried out in this work. It would require a combination of the temporal approximation properties of the θ-methods with the approximation properties of the LSFEM.

To get a strongly coupled formulation for the FSI problem, the formulations for the fluid and the structural part have to be coupled by an appropriate procedure. While several works exist which consider the coupling of formulations based on the Galerkin variational idea, the literature is rare for the LSFEM. In this work one method has been developed to couple a Galerkin formulation for the structural part with a least squares fluid formulation. The ideas used therein are mainly adaptations of ideas proposed for the Galerkin methods.

The second coupling method has already been proposed in literature. This is not really astonishing as it is the consistent way of imposing such coupling conditions within the least squares framework. An analysis of the stability of this coupled system is still missing. The selection of the norms for the residual of the coupling conditions was based on heuristic arguments and clearly has to be mathematically verified. But similar to the fluid part, computational experiments were performed to analyse the general numerical behaviour of the formulation. For a simple test case it was possible to demonstrate that the formulation does work as expected.

Finally the pure least squares formulation and the LSFEM-Galerkin formulation were tested with a more complicated problem. Although the used problem is still far away from practially relevant problems, it exhibits a highly instationary solution and furthermore has the advantage of being solved by several other authors as well. Especially the LSFEM-Galerkin formulation was able to reproduce the results with high accuracy. Unfortunately this formulation currently does not give any real benefits regarding the efficient solution of the resulting system of equations. The pure least squares formulation gave reasonable results when the spectrum was considered. Currently, the reasons for the differences in the time series are unclear, but it seems generally possible to obtain a working pure least squares formulation.

Now the discussion of the initially posed question wheter the LSFEM offers real advantages over the established methods for FSI problems remains. A final answer clearly is not possible as the LSFEM is still an active area of research. The reader of Jiangs book about the LSFEM [83] gets the impression that the LSFEM is the universal tool for partial differential equations and can be used in a kind of black box manner. This impression is definitively wrong.

Due to the reformulation of the original problem the LSFEM gains a lot of stability which makes it easy to solve a driven cavity for example within an afternoon of work. But

as soon as accurate solutions are required the situation completely changes. For many problems the standard L^2 LSFEM does not yield bilinear forms which are coercive in the appropriate norms. Therefore non standard approaches like the negative norm LSFEM, the weighted LSFEM or the augmented LSFEM have to be used to achieve theoretically optimal convergence results. Some of these approaches are even more difficult to implement than the standard stabilisation techniques for the Galerkin FEM, while others like the augmented LSFEM suffer from high regularity demands on the solution and a lower accuracy on a given triangulation.

Another drawback of the LSFEM is the lack of conservation. If this problem can be overcome by some new ideas is not clear yet. Despite this theoretical disadvantage, the numerical results presented in this work indicate that the LSFEM is capable of producing correct solutions for FSI problems.

Hence the LSFEM could still be an interesting alternative if it is possible to find efficient solution techniques. Several articles propose efficient preconditioners for the LSFEM. If these do work well in the context of FSI problems has not been analysed in this work.

8.2 Outlook

In the previous section already some open questions in the LSFEM were addressed. These will be discussed in more detail and be extended by some other interesting issues in this section.

As it was already mentioned, the analysis of the LSFEM for transient problems is still in its beginnings. While in this work some first stability estimates were provided, the extension to a full error analysis would be clearly interesting and important. Also other least squares formulations like the one for the Stokes equations for transient flows could be analysed with these methods. This task should be comparatively straightforward, while an analysis of the least squares formulation for the Navier-Stokes equations for transient flows seems to be quite challenging due to the nonlinearity.

For the equations of elasticity the situation is similar. While the linear equations for stationary problems have been thoroughly analysed in literature and the linear equations of elastodynamics have been analysed in this work, no mathematical results are available for least squares formulations of the nonlinear equations. Especially the introduction of the geometric nonlinearities would be a great step forward to a more realistic model.

Another interesting subproblem which has been left open in this work is the stability of the least squares formulation for the fully coupled problem. In this work, some assumptions were imposed on the solutions on the interface to "secure" the formulation. Clearly this is not completely satisfying from a mathematical point of view. If it is possible to establish coercivity estimates for the resulting bilinear form, these would automatically imply existence and uniqueness of a solution. Hence the analysis could

also give some insights of general interest. Considering the nonlinear problem, similar results could probably be established although it will be more difficult.

Beside these more theoretical problems, some interesting variations of the LSFEM have not yet been tested. As the discussion about the lack of mass conservation revealed, the LSFEM sometimes completely fails to produce reasonable results due to the lack of local and global mass conservation. Using discretely divergence free elements would clearly circumvent this problem, while the beneficial properties of the LSFEM would remain unchanged. To our knowledge this approach has not yet been tried out. Alternatively the use of Raviart-Thomas spaces for the velocity field could offer advantages regarding the mass conservation. These elements can represent pointwise divergence free fields. Therefore it seems reasonable to assume that the minimum of the functional would coincide with a divergence free velocity field. Due to the lack of continuity of the Raviart Thomas spaces, the standard L^2 LSFEM seems to be not suited for this approach. But in conjunction with the negative norm methods is might be possible to find a working formulation.

A recent article from Cai et.al. [41] demonstrated the use of a formulation derived from the $v\sigma$-formulation for linear elasticity for the Stokes equations. First it should be noted that this formulation could easily be extended to transient problems by the same method which was used for the equations of linear elastodynamics in this work. The physical stresses appear as problem unknowns within the formulation and are approximated by Raviart-Thomas spaces. Hence if this formulation is coupled with the formulation proposed for linear elasticity in this work, the equivalence of the normal tractions is automatically satisfied by the Raviart-Thomas spaces. The same holds for the velocity space. Therefore a strongly coupled least squares formulation for a FSI problem can be implemented on a single triangulation by simply changing the material parameters on the different elements representing either the fluid or the structure.

Regarding efficient solution algorithms for the coupled problem, especially multigrid algorithms are very promising. But currently only a few articles have considered the efficient solution of the Navier-Stokes equations for high Reynolds numbers. When the convection becomes dominant, the constants in the local ellipticity estimates change and hence would require modified multigrid algorithms. This problem is currently unresolved for the Navier-Stokes equations.

A Results from Literature

A.1 Coercivity Result for the Laplace Equation

.
For the bilinear form which stems from the LSFEM functional for the diffusion equation:

$$\mathcal{B}(\boldsymbol{u},\boldsymbol{v}) = \left(p - \frac{\partial u}{\partial x}, q - \frac{\partial v}{\partial x}\right)_{0,\Omega} + \left(\frac{\partial p}{\partial x}, \frac{\partial q}{\partial x}\right)_{0,\Omega} \tag{A.1}$$

with $\boldsymbol{u} = (u,p)^T$ and $\boldsymbol{v} = (v,q)^T$, the following theorem from [83] shows full H^1-coercivity:

Theorem 18 *For all $(u,p) \in H_0^1(\Omega) \times H^1(\Omega)$ exists a positive constant c_1 independent of (u,p) such that:*

$$\|u\|_{1,\Omega}^2 + \|p\|_{1,\Omega}^2 \leq c_1 \mathcal{B}((u,p),(u,p))$$

Furthermore there exists a positive constant c_2 independent of (u,p) such that:

$$\mathcal{B}((u,p),(v,q)) \leq c_2(\|u\|_{1,\Omega}\|v\|_{1,\Omega} + \|p\|_{1,\Omega}\|q\|_{1,\Omega}) \quad \forall (u,p),(v,q) \in H_0^1(\Omega) \times H^1(\Omega)$$

Proof:(From [83]) The upper bound can easily be verified by the Cauchy-Schwarz and the triangle inequality.
From the definition of the bilinear form it is clear that

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) \geq \left\| \frac{\partial q}{\partial x} \right\|_{0,\Omega}^2 \tag{A.2}$$

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) \geq \left\| q - \frac{\partial v}{\partial x} \right\|_{0,\Omega}^2 \tag{A.3}$$

For the lower bound a first estimate

$$\begin{aligned} \mathcal{B}(\boldsymbol{v},\boldsymbol{v}) &= \left\| \frac{\partial q}{\partial x} \right\|_{0,\Omega}^2 + \|q\|_{0,\Omega}^2 + \left\| \frac{\partial v}{\partial x} \right\|_{0,\Omega}^2 + -2\left(q, \frac{\partial v}{\partial x}\right)_{0,\Omega} \\ &\geq \left\| \frac{\partial q}{\partial x} \right\|_{0,\Omega}^2 + \|q\|_{0,\Omega}^2 + \|v\|_{0,\Omega}^2 + 2\left(\frac{\partial q}{\partial x}, v\right)_{0,\Omega} \end{aligned} \tag{A.4}$$

can be obtained by partial integration, the Friedrichs inequality and the boundary conditions. Thus

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) \geq \left\| \frac{\partial q}{\partial x} + v \right\|_{0,\Omega}^2 + \|q\|_{0,\Omega}^2 \tag{A.5}$$

and

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) \geq \left\| \frac{\partial q}{\partial x} + v \right\|_{0,\Omega}^2 \tag{A.6}$$

$$\mathcal{B}(\boldsymbol{v},\boldsymbol{v}) \geq \|q\|_{0,\Omega}^2 \tag{A.7}$$

The combination of Eq. (A.2) and Eq. (A.6) gives:

$$2(\mathcal{B}(\boldsymbol{v},\boldsymbol{v}))^{1/2} \geq \left\| \frac{\partial q}{\partial x} \right\|_{0,\Omega} + \left\| \frac{\partial q}{\partial x} + v \right\|_{0,\Omega} = \left\| -\frac{\partial q}{\partial x} \right\|_{0,\Omega} + \left\| \frac{\partial q}{\partial x} + v \right\|_{0,\Omega} \geq \|v\|_{0,\Omega} \tag{A.8}$$

Similarly Eq. (A.3) and Eq. (A.7) lead to:

$$2(\mathcal{B}(\boldsymbol{v},\boldsymbol{v}))^{1/2} \geq \left\| q - \frac{\partial v}{\partial x} \right\|_{0,\Omega} + \|q\|_{0,\Omega} = \left\| \frac{\partial v}{\partial x} - q \right\|_{0,\Omega} + \|q\|_{0,\Omega} \geq \left\| \frac{\partial v}{\partial x} \right\|_{0,\Omega} \tag{A.9}$$

Eq. (A.2), Eq. (A.7), Eq. (A.8) and Eq. (A.9) show then the lower bound □.

A.2 Galerkin Structural Formulation

In 2D the stiffness matrix $\mathbf{K}$ is computed by:

$$\mathbf{K} = \int_\Omega \mathbf{B}(\boldsymbol{w})^T \mathbf{D}\mathbf{B}(\boldsymbol{u})\, d\Omega \tag{A.10}$$

where in 2D

$$\mathbf{B}(\boldsymbol{w}) = \begin{pmatrix} \frac{\partial w_x}{\partial x} & 0 \\ 0 & \frac{\partial w_y}{\partial y} \\ \frac{\partial w_x}{\partial y} & \frac{\partial w_x}{\partial y} \end{pmatrix} \tag{A.11}$$

and

$$\mathbf{D} = \frac{E}{1-\nu^2} \begin{pmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{pmatrix}, \quad \mathbf{D} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & \frac{1-2\nu}{2} \end{pmatrix} \tag{A.12}$$

for the plane stress formulation and the plane strain formulation respectively. The mass matrix is defined by

$$\mathbf{M} = \int_\Omega wu \, \mathrm{d}\Omega \tag{A.13}$$

B Least Squares Functionals

In this section the functionals, for the different least squares formulations, which are used in this work are summarized.

B.1 Stokes $up\omega$-formulation

Equations:

$$\begin{aligned} \nu\nabla\times\omega+\nabla p &= \boldsymbol{f}_1 \quad \text{in} \quad \Omega && \text{(B.1)}\\ \nabla\cdot\boldsymbol{v} &= 0 \quad \text{in} \quad \Omega && \text{(B.2)}\\ \omega-\nabla\times\boldsymbol{v} &= 0 \quad \text{in} \quad \Omega && \text{(B.3)} \end{aligned}$$

B.1.1 Weighted L_2 formulation

Least squares functional:

$$\mathcal{J}(\boldsymbol{v},\omega,p)=\frac{1}{2}(\alpha_0^F\|\nu\nabla\times\omega+\nabla p-\boldsymbol{f}_1\|_{0,\Omega}^2+\alpha_1^F\|\nabla\cdot\boldsymbol{v}\|_{0,\Omega}^2+\alpha_2^F\|\omega-\nabla\times\boldsymbol{v}\|_{0,\Omega}^2) \quad \text{(B.4)}$$

Bilinear form:

$$\begin{aligned} \mathcal{B}((\boldsymbol{v},\omega,p),(\boldsymbol{w},\kappa,q)) =& \alpha_0^F(\nu\nabla\times\omega+\nabla p,\nu\nabla\times\kappa+\nabla q)_{0,\Omega}+\\ & \alpha_1^F(\nabla\cdot\boldsymbol{v},\nabla\cdot\boldsymbol{w})_{0,\Omega}+\\ & \alpha_2^F(\omega-\nabla\times\boldsymbol{v},\kappa-\nabla\times\boldsymbol{w})_{0,\Omega} \end{aligned} \quad \text{(B.5)}$$

B.1.2 Negative norm formulation

Least squares functional:

$$\mathcal{J}(\boldsymbol{v},\omega,p)=\frac{1}{2}(\alpha_0^F\|\nu\nabla\times\omega+\nabla p-\boldsymbol{f}_1\|_{-1,\Omega}^2+\alpha_1^F\|\nabla\cdot\boldsymbol{v}\|_{0,\Omega}^2+\alpha_2^F\|\omega-\nabla\times\boldsymbol{v}\|_{0,\Omega}^2) \quad \text{(B.6)}$$

Bilinear form:

$$\begin{aligned}\mathcal{B}((\boldsymbol{v},\omega,p),(\boldsymbol{w},\kappa,q)) =& \alpha_0^F(\mathcal{S}(\nu\nabla\times\omega+\nabla p),\nu\nabla\times\kappa+\nabla q)_{0,\Omega}+\\ & \alpha_1^F(\nabla\cdot\boldsymbol{v},\nabla\cdot\boldsymbol{w})_{0,\Omega}+\\ & \alpha_2^F(\omega-\nabla\times\boldsymbol{v},\kappa-\nabla\times\boldsymbol{w})_{0,\Omega}\end{aligned} \tag{B.7}$$

B.2 Stokes *uUp*-formulation

Equations:

$$\nu\nabla\cdot\mathbf{L}+\nabla p = \boldsymbol{f}_1 \quad \text{in} \quad \Omega \tag{B.8}$$

$$\nabla\cdot\boldsymbol{v} = 0 \quad \text{in} \quad \Omega \tag{B.9}$$

$$\mathbf{L}-\nabla\boldsymbol{v} = 0 \quad \text{in} \quad \Omega \tag{B.10}$$

B.2.1 Weighted L_2 formulation

Least squares functional:

$$\mathcal{J}(\boldsymbol{v},\mathbf{L},p)=\frac{1}{2}(\alpha_0^F\|-\nu\nabla\cdot\mathbf{L}+\nabla p\|_{0,\Omega}^2+\alpha_1^F\|\nabla\cdot\boldsymbol{v}\|_{0,\Omega}^2+\alpha_2^F\|\mathbf{L}-\nabla\boldsymbol{v}\|_{0,\Omega}^2) \tag{B.11}$$

Bilinear form:

$$\begin{aligned}\mathcal{B}((\boldsymbol{v},\mathbf{L},p),(\boldsymbol{w},\mathbf{M},q)) =& \alpha_0^F(-\nu(\nabla^T\mathbf{L})^T+\nabla p,-\nu(\nabla^T\mathbf{M})^T+\nabla q)_{0,\Omega}+\\ & \alpha_1^F(\nabla\cdot\boldsymbol{v},\nabla\cdot\boldsymbol{w})_{0,\Omega}+\\ & \alpha_2^F(\mathbf{L}-\nabla\boldsymbol{v},\mathbf{M}-\nabla\boldsymbol{w})_{0,\Omega}\end{aligned} \tag{B.12}$$

B.2.2 Negative norm formulation

Least squares functional:

$$\mathcal{J}(\boldsymbol{v},\mathbf{L},p)=\frac{1}{2}(\alpha_0^F\|-\nu\nabla\cdot\mathbf{L}+\nabla p\|_{-1,\Omega}^2+\alpha_1^F\|\nabla\cdot\boldsymbol{v}\|_{0,\Omega}^2+\alpha_2^F\|\mathbf{L}-\nabla\boldsymbol{v}\|_{0,\Omega}^2) \tag{B.13}$$

Bilinear form:

$$\begin{aligned}\mathcal{B}((\boldsymbol{v},\mathbf{L},p),(\boldsymbol{w},\mathbf{M},q)) =& \alpha_0^F(\mathcal{S}(-\nu(\nabla^T\mathbf{L})^T+\nabla p),-\nu(\nabla^T\mathbf{M})^T+\nabla q)_{\Omega}+\\ & \alpha_1^F(\nabla\cdot\boldsymbol{v},\nabla\cdot\boldsymbol{w})_{0,\Omega}+\\ & \alpha_2^F(\mathbf{L}-\nabla\boldsymbol{v},\mathbf{M}-\nabla\boldsymbol{w})_{0,\Omega}.\end{aligned} \tag{B.14}$$

B.3 Navier-Stokes $up\omega$-formulation

Linearised equations:

$$\begin{aligned} \boldsymbol{v}^{k+1}\nabla\boldsymbol{v}^k + \boldsymbol{v}^k\nabla\boldsymbol{v}^{k+1} + \nu\nabla\times\omega^{k+1} + \nabla p^{k+1} &= \boldsymbol{f}_1 + \boldsymbol{v}^k\nabla\boldsymbol{v}^k \quad \text{in} \quad \Omega & \text{(B.15)}\\ \nabla\cdot\boldsymbol{v}^{k+1} &= 0 \quad \text{in} \quad \Omega & \text{(B.16)}\\ \omega^{k+1} - \nabla\times\boldsymbol{v}^{k=1} &= 0 \quad \text{in} \quad \Omega & \text{(B.17)} \end{aligned}$$

Least squares functional:

$$\begin{aligned} &\mathcal{J}(\boldsymbol{v}^{k+1},\omega^{k+1},p^{k+1}) = \\ &\quad \frac{1}{2}(\alpha_0^F\|\boldsymbol{v}^{k+1}\nabla\boldsymbol{v}^k + \boldsymbol{v}^k\nabla\boldsymbol{v}^{k+1}\nu\nabla\times\omega^{k+1} + \nabla p^{k+1} - \tilde{\boldsymbol{f}}_1\|_{0,\Omega}^2 + \\ &\quad\quad \alpha_1^F\|\nabla\cdot\boldsymbol{v}^{k+1}\|_{0,\Omega}^2 + \alpha_2^F\|\omega^{k+1} - \nabla\times\boldsymbol{v}^{k+1}\|_{0,\Omega}^2) \end{aligned} \tag{B.18}$$

with $\tilde{\boldsymbol{f}}_1 = \boldsymbol{f}_1 + \boldsymbol{v}^k\nabla\boldsymbol{v}^k$
Bilinear form:

$$\begin{aligned} \mathcal{B}((\boldsymbol{v}^{k+1},\omega^{k+1},p^{k+1}),(\boldsymbol{w},\kappa,q)) = \alpha_0^F(&\boldsymbol{v}^{k+1}\nabla\boldsymbol{v}^k + \boldsymbol{v}^k\nabla\boldsymbol{v}^{k+1} + \nu\nabla\times\omega^{k+1} + \nabla p^{k+1},\\ &\boldsymbol{v}^{k+1}\nabla\boldsymbol{w} + \boldsymbol{w}^k\nabla\boldsymbol{v}^{k+1} + \nu\nabla\times\kappa + \nabla q)_{0,\Omega} +\\ &\alpha_1^F(\nabla\cdot\boldsymbol{v}^{k+1},\nabla\cdot\boldsymbol{w})_{0,\Omega} +\\ &\alpha_2^F(\omega^{k+1} - \nabla\times\boldsymbol{v}^{k+1},\kappa - \nabla\times\boldsymbol{w})_{0,\Omega} \end{aligned} \tag{B.19}$$

B.4 Navier-Stokes uUp-formulation

Linearised equations:

$$\begin{aligned} \boldsymbol{v}^{k+1}\mathbf{L}^k + \boldsymbol{v}^k\mathbf{L}^{k+1} + \nu\nabla\cdot\mathbf{L}^{k+1} + \nabla p^{k+1} &= \boldsymbol{f}_1 + \boldsymbol{v}^k\mathbf{L}^k \quad \text{in} \quad \Omega & \text{(B.20)}\\ \nabla\cdot\boldsymbol{v}^{k+1} &= 0 \quad \text{in} \quad \Omega & \text{(B.21)}\\ \mathbf{L}^{k+1} - \nabla\boldsymbol{v}^{k+1} &= 0 \quad \text{in} \quad \Omega & \text{(B.22)} \end{aligned}$$

Least squares functional:

$$\begin{aligned} &\mathcal{J}(\boldsymbol{v}^{k+1},\mathbf{L}^{k+1},p^{k+1}) = \\ &\quad \frac{1}{2}(\alpha_0^F\|\boldsymbol{v}^{k+1}\mathbf{L}^k + \boldsymbol{v}^k\mathbf{L}^{k+1} + \nu\nabla\cdot\mathbf{L}^{k+1} + \nabla p^{k+1} - \tilde{\boldsymbol{f}}_1\|_{0,\Omega}^2 + \\ &\quad\quad \alpha_1^F\|\nabla\cdot\boldsymbol{v}\|_{0,\Omega}^2 + \alpha_2^F\|\mathbf{L} - \nabla\boldsymbol{v}\|_{0,\Omega}^2) \end{aligned} \tag{B.23}$$

with $\tilde{\boldsymbol{f}}_1 = \boldsymbol{f}_1 + \boldsymbol{v}^k\mathbf{L}^k$.

Bilinear form:

$$\begin{aligned}\mathcal{B}((\boldsymbol{v}^{k+1},\mathbf{L}^{k+1},p^{k+1}),(\boldsymbol{w},\mathbf{M},q)) = \\ \alpha_0^F(\boldsymbol{v}^{k+1}\mathbf{L}^k + \boldsymbol{v}^k\mathbf{L}^{k+1} + \nu\nabla\cdot\mathbf{L}^{k+1} + \nabla p^{k+1}, \\ \boldsymbol{w}\mathbf{L}^k + \boldsymbol{v}^k\boldsymbol{M} + \nu\nabla\cdot\boldsymbol{M} + \nabla q)_{0,\Omega} + \\ \alpha_1^F(\nabla\cdot\boldsymbol{v},\nabla\cdot\boldsymbol{w})_{0,\Omega} + \\ \alpha_2^F(\mathbf{L}-\nabla\boldsymbol{v},\mathbf{M}-\nabla\boldsymbol{w})_{0,\Omega}\end{aligned} \tag{B.24}$$

B.5 ALE Navier-Stokes *upω*-formulation for transient flows

Linearised and time discretised equations:

$$\frac{\boldsymbol{v}_{n+1}^{k+1}}{\Delta t} + \theta(\boldsymbol{v}_{n+1}^{k+1}\cdot\nabla\boldsymbol{v}_{n+1}^{k} + \boldsymbol{v}_{n+1}^{k}\cdot\nabla\boldsymbol{v}_{n+1}^{k+1} - \boldsymbol{v}_{\Phi_{n+1}}\cdot\nabla\boldsymbol{v}_{n+1}^{k+1} + \tag{B.25}$$

$$\nu_f\nabla\times\omega_{n+1}^{k+1}) + \nabla p = \tag{B.26}$$

$$\frac{\boldsymbol{v}_{n+1}^{k+1}}{\Delta t} - (1-\theta)((\boldsymbol{v}_n - \boldsymbol{v}_{\Phi_n})\cdot\nabla\boldsymbol{v}_n + \nu_f\nabla\times\omega_n + \boldsymbol{f}_n) + \tag{B.27}$$

$$\theta(\boldsymbol{v}_{n+1}^{k}\nabla\cdot\boldsymbol{v}_{n+1}^{k} + \boldsymbol{f}_{n+1}) \quad \text{in} \quad \Omega \tag{B.28}$$

$$\nabla\cdot\boldsymbol{v}_{n+1}^{k+1} = 0 \quad \text{in} \quad \Omega \tag{B.29}$$

$$\omega_{n+1}^{k+1} - \nabla\times\boldsymbol{v}_{n+1}^{k+1} = 0 \quad \text{in} \quad \Omega \tag{B.30}$$

Least squares functional:

$$\begin{aligned}\mathcal{J}(\boldsymbol{v}_{n+1}^{k+1},&\omega_{n+1}^{k+1},p_{n+1}^{k+1}) = \\ &\frac{1}{2}(\alpha_0^F\|(\boldsymbol{v}_{n+1}^{k+1}/\Delta t) + \theta(\boldsymbol{v}_{n+1}^{k+1}\cdot\nabla\boldsymbol{v}_{n+1}^{k} + \boldsymbol{v}_{n+1}^{k}\cdot\nabla\boldsymbol{v}_{n+1}^{k+1} - \boldsymbol{v}_{\Phi_{n+1}}\cdot\nabla\boldsymbol{v}_{n+1}^{k+1} + \\ &\qquad \nu_f\nabla\times\omega_{n+1}^{k+1}) + \nabla p - \tilde{\boldsymbol{f}}_1\|_{0,\Omega}^2 + \\ &\alpha_1^F\|\nabla\cdot\boldsymbol{v}_{n+1}^{k+1}\|_{0,\Omega}^2 + \alpha_2^F\|\omega_{n+1}^{k+1} - \nabla\times\boldsymbol{v}_{n+1}^{k+1}\|_{0,\Omega}^2)\end{aligned} \tag{B.31}$$

with $\tilde{\boldsymbol{f}}_1 = (\boldsymbol{v}_{n+1}^{k+1}/\Delta t) - (1-\theta)((\boldsymbol{v}_n - \boldsymbol{v}_{\Phi_n})\cdot\nabla\boldsymbol{v}_n + \nu_f\nabla\times\omega_n + \boldsymbol{f}_n) + \theta(\boldsymbol{v}_{n+1}^{k}\nabla\cdot\boldsymbol{v}_{n+1}^{k} + \boldsymbol{f}_{n+1})$

Bilinear form:

$$\begin{aligned}
&\mathcal{B}((\boldsymbol{v}_{n+1}^{k+1},\omega_{n+1}^{k+1},p_{n+1}^{k+1}),(\boldsymbol{w},\kappa,q)) = \\
&\quad \alpha_0^F((\boldsymbol{v}_{n+1}^{k+1}/\Delta t)+\theta(\boldsymbol{v}_{n+1}^{k+1}\cdot\nabla\boldsymbol{v}_{n+1}^{k}+\boldsymbol{v}_{n+1}^{k}\cdot\nabla\boldsymbol{v}_{n+1}^{k+1}-\boldsymbol{v}_{\Phi_{n+1}}\cdot\nabla\boldsymbol{v}_{n+1}^{k+1}+ \\
&\qquad \nu_f\nabla\times\omega_{n+1}^{k+1})+\nabla p, \\
&\qquad (\boldsymbol{w}/\Delta t)+\theta(\boldsymbol{w}\cdot\nabla\boldsymbol{v}_{n+1}^{k}+\boldsymbol{v}_{n+1}^{k}\cdot\nabla\boldsymbol{w}-\boldsymbol{v}_{\Phi_{n+1}}\cdot\nabla\boldsymbol{w}+ \\
&\qquad \nu_f\nabla\times\kappa)+\nabla q)_{0,\Omega}+ \\
&\quad \alpha_1^F(\nabla\cdot\boldsymbol{v}^{k+1},\nabla\cdot\boldsymbol{w})_{0,\Omega}+ \\
&\quad \alpha_2^F(\omega^{k+1}-\nabla\times\boldsymbol{v}^{k+1},\kappa-\nabla\times\boldsymbol{w})_{0,\Omega}
\end{aligned} \tag{B.32}$$

B.6 Navier-Stokes uUp-formulation for transient flows

Linearised and time discretised equations:

$$\frac{\boldsymbol{v}_{n+1}^{k+1}}{\Delta t}+\theta(\boldsymbol{v}_{n+1}^{k+1}\mathbf{L}_{n+1}^{k}+\boldsymbol{v}_{n+1}^{k}\mathbf{L}_{n+1}^{k+1}-\nu_f\nabla\cdot\mathbf{L}_{n+1}^{k+1})+\nabla p_{n+1}^{k+1}= \tag{B.33}$$

$$\frac{\boldsymbol{v}_n}{\Delta t}-(1-\theta)(\boldsymbol{v}_n\mathbf{L}_n-\nu_f\nabla\cdot\mathbf{L}_{n+1}^{k+1}+\boldsymbol{f}_n)+ \tag{B.34}$$

$$\theta(\boldsymbol{v}_{n+1}^{k}\mathbf{L}_{n+1}^{k}+\boldsymbol{f}_{n+1}) \quad \text{in} \quad \Omega \tag{B.35}$$

$$\nabla\cdot\boldsymbol{v}_{n+1}^{k+1}=0 \quad \text{in} \quad \Omega \tag{B.36}$$

$$\mathbf{L}_{n+1}^{k+1}-\nabla\boldsymbol{v}_{n+1}^{k+1}=0 \quad \text{in} \quad \Omega \tag{B.37}$$

Least squares functional:

$$\begin{aligned}
&\mathcal{J}(\boldsymbol{v}_{n+1}^{k+1},\mathbf{L}_{n+1}^{k+1},p_{n+1}^{k+1}) = \\
&\qquad \frac{1}{2}(\alpha_0^F\|(\boldsymbol{v}_{n+1}^{k+1}/\Delta t)+\theta(\boldsymbol{v}_{n+1}^{k+1}\mathbf{L}_{n+1}^{k}+\boldsymbol{v}_{n+1}^{k}\mathbf{L}_{n+1}^{k+1} \\
&\qquad\qquad -\nu_f\nabla\cdot\mathbf{L}_{n+1}^{k+1})+\nabla p_{n+1}^{k+1}-\tilde{\boldsymbol{f}}_1\|_{0,\Omega}^2+ \\
&\qquad\quad \alpha_1^F\|\nabla\cdot\boldsymbol{v}_{n+1}^{k+1}\|_{0,\Omega}^2+\alpha_2^F\|\mathbf{L}_{n+1}^{k+1}-\nabla\boldsymbol{v}_{n+1}^{k+1}\|_{0,\Omega}^2)
\end{aligned} \tag{B.38}$$

with $\tilde{\boldsymbol{f}}_1=(\boldsymbol{v}_n/\Delta t)-(1-\theta)(\boldsymbol{v}_n\mathbf{L}_n-\nu_f\nabla\cdot\mathbf{L}_{n+1}^{k+1}+\boldsymbol{f}_n)+\theta(\boldsymbol{v}_{n+1}^{k}\mathbf{L}_{n+1}^{k}+\boldsymbol{f}_{n+1})$
Bilinear form:

$$\begin{aligned}
&\mathcal{B}((\boldsymbol{v}_{n+1}^{k+1},\mathbf{L}_{n+1}^{k+1},p_{n+1}^{k+1}),(\boldsymbol{w},\mathbf{M},q)) = \\
&\quad \alpha_0^F((\boldsymbol{v}_{n+1}^{k+1}/\Delta t)+\theta(\boldsymbol{v}_{n+1}^{k+1}\mathbf{L}_{n+1}^{k}+\boldsymbol{v}_{n+1}^{k}\mathbf{L}_{n+1}^{k+1}-\nu_f\nabla\cdot\mathbf{L}_{n+1}^{k+1})+\nabla p_{n+1}^{k+1}, \\
&\qquad (\boldsymbol{w}/\Delta t)+\theta(\boldsymbol{w}\mathbf{L}_{n+1}^{k}+\boldsymbol{v}_{n+1}^{k}\boldsymbol{M}-\nu_f\nabla\cdot\boldsymbol{M}+\nabla q)_{0,\Omega}+ \\
&\quad \alpha_1^F(\nabla\cdot\boldsymbol{v}_{n+1}^{k+1},\nabla\cdot\boldsymbol{w})_{0,\Omega}+ \\
&\quad \alpha_2^F(\mathbf{L}_{n+1}^{k+1}-\nabla\boldsymbol{v}_{n+1}^{k+1},\mathbf{M}-\nabla\boldsymbol{w})_{0,\Omega}
\end{aligned} \tag{B.39}$$

B.7 Linear Elasticity $u\sigma$-formulation

B.7.1 Unconstrained

Equations:

$$\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon = 0 \quad \text{in} \quad \Omega \tag{B.40}$$

$$\nabla \cdot \sigma = \boldsymbol{f} \quad \text{in} \quad \Omega \tag{B.41}$$

Least squares functional:

$$\mathcal{J}(\sigma, \boldsymbol{v}) = \frac{1}{2}(\alpha_0^S \|\nabla \cdot \sigma - \boldsymbol{f}\|_{0,\Omega}^2 + \alpha_1^S \|\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon\|_{0,\Omega}^2) \tag{B.42}$$

Bilinear form:

$$\begin{aligned}\mathcal{B}((\sigma, \boldsymbol{v}), (\tau, \boldsymbol{w})) =& \alpha_0^S(\nabla \cdot \sigma, \nabla \cdot \tau)_{0,\Omega} + \\ & \alpha_1^S(\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}), \mathbb{C}^{-1/2}\tau - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{w}))_{0,\Omega}\end{aligned} \tag{B.43}$$

B.7.2 With Symmetry Constraint

Equations:

$$\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon = 0 \quad \text{in} \quad \Omega \tag{B.44}$$

$$\nabla \cdot \sigma = 0 \quad \text{in} \quad \Omega \tag{B.45}$$

$$\sigma^T - \sigma = 0 \quad \text{in} \quad \Omega \tag{B.46}$$

Least squares functional:

$$\mathcal{J}(\sigma, \boldsymbol{v}) = \frac{1}{2}(\alpha_0^S \|\nabla \cdot \sigma - \boldsymbol{f}\|_{0,\Omega}^2 + \alpha_1^S \|\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon\|_{0,\Omega}^2 + \alpha_2^S \|\sigma - \sigma^T\|_{0,\Omega}^2) \tag{B.47}$$

Bilinear form:

$$\begin{aligned}\mathcal{B}((\sigma, \boldsymbol{v}), (\tau, \boldsymbol{w})) =& \alpha_0^S(\nabla \cdot \sigma, \nabla \cdot \tau)_{0,\Omega} + \\ & \alpha_1^S(\mathbb{C}^{-1/2}\sigma - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}), \mathbb{C}^{-1/2}\tau - \mathbb{C}^{1/2}\varepsilon(\boldsymbol{w}))_{0,\Omega} + \\ & \alpha_2^S(\sigma - \sigma^T, \tau - \tau^T)_{0,\Omega}\end{aligned} \tag{B.48}$$

B.8 Linear Elastodynamics $v\boldsymbol{H}$-formulation

Time discretised equations:

$$\frac{\mathbf{H}_{n+1}}{\Delta t} - \theta\nabla \boldsymbol{v}_{n+1} = \frac{\mathbf{H}_n}{\Delta t} + (1-\theta)\nabla \boldsymbol{v}_n \quad \text{in} \quad \Omega \tag{B.49}$$

$$\frac{\boldsymbol{v}_{n+1}}{\Delta t} - \theta\nabla \cdot \mathbb{A}\mathbf{H}_{n+1} = \frac{\boldsymbol{v}_n}{\Delta t} + (1-\theta)\nabla \cdot \mathbb{A}\mathbf{H}_n \quad \text{in} \quad \Omega \tag{B.50}$$

$$\frac{\nabla \times \mathbf{H}_{n+1}}{\Delta t} = \frac{\nabla \times \mathbf{H}_n}{\Delta t} \quad \text{in} \quad \Omega \tag{B.51}$$

with

$$\mathbb{A} = \begin{pmatrix} c_1 & 0 & 0 & c_3 \\ 0 & c_2 & c_2 & 0 \\ 0 & c_2 & c_2 & 0 \\ c_3 & 0 & 0 & c_1 \end{pmatrix} \tag{B.52}$$

and $c_1 = (\lambda + 2\mu)/\rho$, $c_2 = \mu/\rho$ and $c_3 = \lambda/\rho$.
Least squares functional:

$$\begin{aligned} \mathcal{J}(\boldsymbol{v}_{n+1}, \mathbf{H}_{n+1}) =& \alpha_0^S \| \frac{\boldsymbol{v}_{n+1}}{\Delta t} - \theta \nabla \cdot \mathbb{A}\mathbf{H}_{n+1} - \boldsymbol{f}_1 \|_{0,\Omega}^2 + \\ & \alpha_1^S \| \frac{\mathbf{H}_{n+1}}{\Delta t} - \theta \nabla \boldsymbol{v}_{n+1} - \boldsymbol{f}_2 \|_{0,\Omega}^2 + \alpha_2^S \| \frac{\nabla \times \mathbf{H}_{n+1}}{\Delta t} - \boldsymbol{f}_3 \|_{0,\Omega}^2 \end{aligned} \tag{B.53}$$

with $\boldsymbol{f}_1 = \frac{\boldsymbol{v}_n}{\Delta t} + (1-\theta)\nabla \cdot \mathbb{A}\mathbf{H}_n$, $\boldsymbol{f}_2 = \frac{\mathbf{H}_n}{\Delta t} + (1-\theta)\nabla \boldsymbol{v}_n$ and $\boldsymbol{f}_3 = \frac{\nabla \times \mathbf{H}_n}{\Delta t}$
Bilinear form:

$$\begin{aligned} \mathcal{B}((\boldsymbol{v}_{n+1}, \mathbf{H}_{n+1}), (\boldsymbol{w}, \boldsymbol{M})) =& \alpha_0^S (\frac{\boldsymbol{v}_{n+1}}{\Delta t} - \theta \nabla \cdot \mathbb{A}\mathbf{H}_{n+1}, \frac{\boldsymbol{w}}{\Delta t} - \theta \nabla \cdot \mathbb{A}\boldsymbol{M})_{0,\Omega} \\ & \alpha_1^S (\frac{\mathbf{H}_{n+1}}{\Delta t} - \theta \nabla \boldsymbol{v}_{n+1}, \frac{\boldsymbol{M}}{\Delta t} - \theta \nabla \boldsymbol{w})_{0,\Omega} \\ & \alpha_2^S (\frac{\nabla \times \mathbf{H}_{n+1}}{\Delta t}, \frac{\nabla \times \boldsymbol{M}}{\Delta t})_{0,\Omega} \end{aligned} \tag{B.54}$$

B.9 Linear Elastodynamics $v\sigma$-formulation

B.9.1 Unconstrained

Equations:

$$\frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}} \sigma_{n+1} - \sqrt{k_1} \mathbb{C}^{1/2} \varepsilon_{n+1} = \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}} \sigma_n + (k_2/\sqrt{k_1}) \mathbb{C}^{1/2} \varepsilon_n \tag{B.55}$$

$$\frac{\sqrt{\rho}}{\sqrt{k_1}} \boldsymbol{v}_{n+1} - \frac{\sqrt{k_1}}{\sqrt{\rho}} \nabla \cdot \sigma_{n+1} = \frac{\sqrt{\rho}}{\sqrt{k_1}} \boldsymbol{v}_n + \frac{k_2}{\sqrt{k_1}\sqrt{\rho}} \nabla \cdot \sigma_n \tag{B.56}$$

with $k_1 = \theta \Delta t$ and $k_2 = (1-\theta)\Delta t$.
Least squares functional:

$$\begin{aligned} \mathcal{J}(\sigma_{n+1}, \boldsymbol{v}_{n+1}) = \frac{1}{2} (& \alpha_0^S \| \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}} \sigma_{n+1} - \sqrt{k_1} \mathbb{C}^{1/2} \varepsilon_{n+1} - \boldsymbol{f}_1 \|_{0,\Omega}^2 + \\ & \alpha_1^S \| \frac{\sqrt{\rho}}{\sqrt{k_1}} \boldsymbol{v}_{n+1} - \frac{\sqrt{k_1}}{\sqrt{\rho}} \nabla \cdot \sigma_{n+1} - \boldsymbol{f}_2 \|_{0,\Omega}^2) \end{aligned} \tag{B.57}$$

with $\boldsymbol{f}_1 = \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}} \sigma_n + (k_2/\sqrt{k_1}) \mathbb{C}^{1/2} \varepsilon_n$, $\boldsymbol{f}_2 = \frac{\sqrt{\rho}}{\sqrt{k_1}} \boldsymbol{v}_n + \frac{k_2}{\sqrt{k_1}\sqrt{\rho}} \nabla \cdot \sigma_n$

Bilinear form:

$$\begin{aligned}\mathcal{B}((\sigma,\boldsymbol{v}),(\tau,\boldsymbol{w})) =& \alpha_0^S(\frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\sigma_{n+1} - \sqrt{k_1}\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}_{n+1}), \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\tau - \sqrt{k_1}\mathbb{C}^{1/2}\varepsilon(\boldsymbol{w}))_{0,\Omega} + \\ & \alpha_1^S(\frac{\sqrt{\rho}}{\sqrt{k_1}}\boldsymbol{v}_{n+1} - \frac{\sqrt{k_1}}{\sqrt{\rho}}\nabla\cdot\sigma_{n+1}, \frac{\sqrt{\rho}}{\sqrt{k_1}}\boldsymbol{w} - \frac{\sqrt{k_1}}{\sqrt{\rho}}\nabla\cdot\tau)_{0,\Omega}\end{aligned} \tag{B.58}$$

B.9.2 With Symmetry Constraint

Equations:

$$\frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\sigma_{n+1} - \sqrt{k_1}\mathbb{C}^{1/2}\varepsilon_{n+1} = \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\sigma_n + (k_2/\sqrt{k_1})\mathbb{C}^{1/2}\varepsilon_n \tag{B.59}$$

$$\frac{\sqrt{\rho}}{\sqrt{k_1}}v_{n+1} - \frac{\sqrt{k_1}}{\sqrt{\rho}}\nabla\cdot\sigma_{n+1} = \frac{\sqrt{\rho}}{\sqrt{k_1}}v_n + \frac{\sqrt{k_2}}{\sqrt{k_1}\sqrt{\rho}}\nabla\cdot\sigma_n \tag{B.60}$$

$$\sigma_{n+1}^T - \sigma_{n+1} = 0 \tag{B.61}$$

Least squares functional:

$$\begin{aligned}\mathcal{J}(\sigma_{n+1},\boldsymbol{v}_{n+1}) = \frac{1}{2}(& \alpha_0^S\|\frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\sigma_{n+1} - \sqrt{k_1}\mathbb{C}^{1/2}\varepsilon_{n+1} - \boldsymbol{f}_1\|_{0,\Omega}^2 + \\ & \alpha_1^S\|\frac{\sqrt{\rho}}{\sqrt{k_1}}\boldsymbol{v}_{n+1} - \frac{\sqrt{k_1}}{\sqrt{\rho}}\nabla\cdot\sigma_{n+1} - \boldsymbol{f}_2\|_{0,\Omega}^2 + \\ & \alpha_2^S\|\sigma - \sigma^T\|_{0,\Omega}^2)\end{aligned} \tag{B.62}$$

with $\boldsymbol{f}_1 = \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\sigma_n + (k_2/\sqrt{k_1})\mathbb{C}^{1/2}\varepsilon_n$, $\boldsymbol{f}_2 = \frac{\sqrt{\rho}}{\sqrt{k_1}}\boldsymbol{v}_n + \frac{k_2}{\sqrt{k_1}\sqrt{\rho}}\nabla\cdot\sigma_n$
Bilinear form:

$$\begin{aligned}\mathcal{B}((\sigma,\boldsymbol{v}),(\tau,\boldsymbol{w})) =& \alpha_0^S(\frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\sigma_{n+1} - \sqrt{k_1}\mathbb{C}^{1/2}\varepsilon(\boldsymbol{v}_{n+1}), \frac{\mathbb{C}^{-1/2}}{\sqrt{k_1}}\tau - \sqrt{k_1}\mathbb{C}^{1/2}\varepsilon(\boldsymbol{w}))_{0,\Omega} + \\ & \alpha_1^S(\frac{\sqrt{\rho}}{\sqrt{k_1}}\boldsymbol{v}_{n+1} - \frac{\sqrt{k_1}}{\sqrt{\rho}}\nabla\cdot\sigma_{n+1}, \frac{\sqrt{\rho}}{\sqrt{k_1}}\boldsymbol{w} - \frac{\sqrt{k_1}}{\sqrt{\rho}}\nabla\cdot\tau)_{0,\Omega} + \\ & \alpha_2^S(\sigma - \sigma^T, \tau - \tau^T)_{0,\Omega}\end{aligned} \tag{B.63}$$

Symbols and Notation

The following table contains a short description of the symbols and notations used in this work. Generally small latin letters denote scalar values and vectors values if they are printed in bold style. Capital bold latin letters are used for matrices and second order tensors. Fourth order tensors are denoted with capital calligraphic letters. If a variable belongs to the structural part it gets the index s. For the fluid part the index f is used.

Mathematical Notation

t	Time
t_0	Start time
T	End time
d	Number of space dimensions
$\mathbb{R}^d$	d-dimensional Euklidian space
Ω	Domain ($\Omega \subset \mathbb{R}^d$)
Γ	Boundary of domain Ω
$L^p(\Omega)$	Space of functions f measurable on Ω such that $\|f\|_{p,\Omega} < \infty$
$W^{k,p}(\Omega)$	Sobolev-Space of functions with weak derivatives in L^p up to order k
$H^k(\Omega)$	Sobolev-Space $W^{k,2}$
$C^q(\Omega)$	Space of continuous function with derivatives up to order of q
$\dot{x}$	Total derivative of x with respect to time
$(x,y)_{\mathbf{V},\Omega}$	scalar product of x and y in a Hilbert space $\mathbf{V}(\Omega)$, over a domain Ω
$\langle x,y\rangle_\Omega$	duality pairing of x and y over a domain Ω
$\|x\|_{\mathbf{V},\Omega}$	Norm of x in a normed space $\mathbf{V}(\Omega)$ over a domain Ω
$\mathbf{X}^T, \boldsymbol{x}^T$	Transponse of a matrix or a vector
α	Multiindex, or some weight for a least squares term
$D^{(\alpha)}$	Weak derivative
$\mathcal{D}\mathcal{F}(u)[v]$	Gâteaux derivative of $\mathcal{F}$
i	Index and also $\sqrt{-1}$

Continuum Mechanics Notation

φ	Mapping from material configuration to current configuration
Ψ	Mapping from material configuration to reference configuration
Φ	Mapping from reference configuration to current configuration
R_X	Material configuration
R_x	Current configuration
R_χ	Reference configuration
$\boldsymbol{X}$	Coordinate vector of point in material configuration
$\boldsymbol{x}$	Coordinate vector of point in current configuration
$\boldsymbol{\chi}$	Coordinate vector of point in reference configuration
$\boldsymbol{u}$	Vector of displacements relative to the material configuration
$\boldsymbol{v}$	Vector of velocities relative to the material configuration
$\boldsymbol{a}$	Vector of accelerations relative to the material configuration
$\boldsymbol{n}$	Outward normal vector
v_x, v_y, v_z	Components of velocity vector $\boldsymbol{v}$
u_x, u_y, u_z	Components of displacement vector $\boldsymbol{u}$
a_x, a_y, a_z	Components of acceleration vector $\boldsymbol{a}$
n_x, n_y, n_y	Components of the normal vector $\boldsymbol{n}$
$\boldsymbol{u}_\Psi$	Displacement relative to the reference configuration
$\boldsymbol{u}_\Phi$	Displacement of reference configuration with respect to the material configuration
$\boldsymbol{v}_\Psi$	Velocity relative to the reference configuration
$\boldsymbol{v}_\Phi$	Velocity of reference configuration with respect to the material configuration
∇_χ	The nabla operator in the reference configuration
∇_x	The nabla operator in the current configuration
∇_X	The nabla operator in the material configuration
f^*	Scalar function in reference configuration
f^{**}	Scalar function in material configuration
$\boldsymbol{n}_C$	Outward normal vector (in current configuration)
$\boldsymbol{n}_M$	Outward normal vector (in material configuration)
$\mathbf{I}$	Identity matrix or Kronecker tensor
$\mathbf{F}$	Deformation gradient
$\dot{\mathbf{F}}$	Material velocity gradient
$\mathbf{L}$	Spatial velocity gradient
$\mathbf{D}$	Symmetric part of $\mathbf{L}$
$\mathbf{W}$	Antisymmetric part of $\mathbf{L}$
$\mathbf{H}$	Displacement gradient
$\mathbf{C}$	Cauchy strain tensor
$\mathbf{E}$	Green Lagrange strain tensor

ε	Linearised Green Lagrange strain tensor
$\dot{\mathbf{E}}$	Material strain rate tensor
$\mathbf{P}$	First Piola Kirchhoff stress tensor
$\mathbf{S}$	Second Piola Kirchhoff stress tensor
σ	Cauchy stress tensor
$\boldsymbol{t}$	Counterclockwise oriented unit tangent vector
$\boldsymbol{t}_C$	Stress vector in current configuration
$\boldsymbol{t}_M$	Stress vector in material configuration

Finite Element Notation

$\mathcal{B}(u, v)$	A bilinear form
$\mathcal{L}$	General differential operator
$\mathcal{R}$	A differential operator representing the boundary conditions
$\mathcal{J}$	General least squares functional
$\mathcal{E}$	Differential operator representing the transient part of a DAE
$\mathcal{C}$	Differential operator representing the constraint of a DAE
$\mathcal{T}$	Trace operator
$\mathcal{T}_h$	Triangulation of the domain Ω
h	Lower bound of the simplex diameters in a triangulation $\mathcal{T}_h$
h_κ	Diameter of a simplex κ
ρ_κ	Diameter of the smallest ball that fits into a simplex κ
σ_κ	Regularity of a simplex κ
$\mathbf{S}_{h,k}(\Omega)$	Standard finite element space of piecewise polynomials of a maximum degree of k on a triangulation $\mathcal{T}_h$
$\mathbf{P}_k$	Space of polynomials up to order k on a d-simplex
$\mathcal{P}_h$	Projection operator
$\mathcal{S}$	Solution operator of the Laplace problem
$\mathcal{S}_h$	Discrete solution operator of the Laplace problem
β	First parameter of the Newmark method
γ	Second parameter of the Newmark method
$\mathbf{K}$	General stiffness matrix, structural stiffness matrix
$\mathbf{M}$	Mass matrix
$\mathcal{E}_l$	Left hand side operator
$\mathcal{E}_r$	Right hand side operator
$\mathcal{I}$	Identity operator
$\boldsymbol{f}, \boldsymbol{f}_i, \boldsymbol{g}, \boldsymbol{w}$	Vector valued function
f, g, f_i	Scalar function
τ	Tensor valued function
$\boldsymbol{p}$	Gradient of a scalar function

p	Usually pressure, sometimes the gradient of a scalar function in 1D
ω	Vorticity
$\mathbf{T}$	Scaled stress tensor
$\tilde{p}$	Kinematic pressure

Equation Parameters

Re	Reynolds number
Re_E	Erguns Reynolds number (porous media benchmark)
D_p	Particle diameter (porous media benchmark)
ε_p	Porosity (porous media benchmark)
H	Channel height (porous media benchmark)
$\bar{\boldsymbol{v}}$	Characteristic velocity
$\bar{L}$	Characteristic length
D	Diameter of cylinder in benchmark problem, Channel diameter
c_D	Drag coefficient
c_L	Lift coefficient
F_D	Horizontal force generated by fluid motion
F_L	Vertical force generated by fluid motion
e_{kin}	Kinetic energy
e_{def}	Internal deformation energy
St	Strouhal number
ρ	Specific mass density
ρ_s	Specific mass density of structural part
ρ_f	Specific mass density of fluid part
η_f	Shear viscosity of the fluid
ν_f	Kinematic viscosity of the fluid
λ	Lamé constant
μ_s	Lamé constant
μ_0	Lower bound for the Lamé parameter μ_s
ν_s	Poisson ratio of the structure
E	Youngs modulus of the structure
c	Wave velocity

Symbols

Γ_D	Dirichlet boundary
Γ_N	Neumann boundary
Γ_{if}	Fluid structure interface
Δt	Time step size

θ	Parameter which determines the properties of the time discretisation
n	Timestep number
k	Maximum polynomial degree of a finite element space
s	Regularity index of a solution u ($u \in H^{s+1}(\Omega)$)
q	General regularity index
c_i	A general constant in estimates
δ	Switch between incompressible Stokes eq. ($\delta = 0$) and linear elasticity ($\delta = 1$)
$\mathbb{A}$	Material tensor
$\mathbb{Q}$	Rotation tensor
$\mathbf{V}$	Rotated displacement gradient ($\mathbf{H} = \mathbb{Q}\mathbf{V}$)
$\mathbb{C}$	Elasticity tensor
$\mathbb{E}$	Compliance tensor ($\mathbb{C}^{-1}$)
k_1	$\theta\Delta t$
k_2	$(1-\theta)\Delta t$
$\boldsymbol{m}$	Stress like unknown for the vT-formulation
$\mathbf{G}$	Material matrix for the vT-formulation
$\mathbf{A}$	Coefficient matrix for ∂x
$\mathbf{B}$	Coefficient matrix for ∂y
$\tilde{\mathbf{B}}$	$\mathbf{A}^{-1}\mathbf{B}$
$\mathbf{C}_1$	Matrix representing a Neumann boundary
$\mathbf{C}_2$	Matrix representing a Dirichlet boundary

Least Squares Weights

α_0^F	Weight for conservation of momentum in fluid part
α_1^F	Weight for incompressibility constraint in fluid part
α_2^F	Weight for constraint defining the new unknowns in fluid part (either the vorticity or the velocity gradient)
α_0^S	Weight for conservation of momentum in structural part
α_1^S	Weight for constraint defining the new unknows in structural part
α_2^S	Weight for stabilising equations in structural part
α_3^S	Weight for weak boundary conditions
α_1^C	Weight for compatibility condition
α_2^C	Weight for equivalence of normal tractions

Bibliography

[1] Agmon, S., A. Douglis and L. Nirenberg: *Estimates Near the Boundary for Solutions of Elliptic Partial Differential Equations Satisfying General Boundary Conditions. I.* Comm. Pure Appl. Math., 12:623–727, 1959.

[2] ANSYS, Inc., 275 Technology Drive, Canonsburg, PA 15371, USA: *ANSYS 6.1 Documentation*, 2002. `http://www.ansys.com`.

[3] Aziz, A.K., R.B. Kellog and A.B. Stephens: *Least squares methods for elliptic systems.* Math. Comp., 44:53–70, 1985.

[4] Bangerth, Wolfgang, Ralf Hartmann and Guido Kanschat: `deal.II` *Differential Equations Analysis Library, Technical Reference.* Interdisciplinary Center for Scientific Computing of the University Heidelberg. `http://www.dealii.org`.

[5] Bathe, Klaus-Jürgen: *Finite Element Procedures.* Prentice-Hall, Englewood Cliffs, N.J., 1996.

[6] Batina, J. T.: *Unsteady Euler airfoil solutions using unstructured dynamic meshes.* AIAA J., 28(8):1381–1388, 1990.

[7] Bell, Brent C. and Karan S. Surana: *p-Version Least Squares Finite Element Formulation for Two-Dimensional, Incompressible, Non-Newtonian Isothermal and Non-Isothermal Fluid Flow.* Internat. J. Numer. Methods Fluids, 18:127–162, 1994.

[8] Bell, Brent C. and Karan S. Surana: *A Space-Time Coupled p-Version Least Squares Finite Element Formulation for Unsteady Two-Dimensional Navier-Stokes Equations.* Internat. J. Numer. Methods Engrg., 39:2593–2618, 1996.

[9] Belytschko, Ted, Wing Kam Liu and Brian Moran: *Nonlinear Finite Elements for Continua and Structures.* John Wiley & Sons, New York, London, Sydney, 2000.

[10] Berndt, M., T. A. Manteuffel and S. F. McCormick: *Local error estimates and adaptive refinement for first-order system least squares (FOSLS).* Elect. Trans. Numer. Anal., 6:35–43, 1997.

[11] Betten, J.: *Finite Elemente für Ingenieure.* Springer-Verlag, Berlin, 1998.

[12] Bochev, P.: *Experiences with negative norm least-square methods for the Navier–Stokes equations.* Elect. Trans. Numer. Anal., 6:44–62, 1997.

[13] Bochev, P., Z. Cai, T. A. Manteuffel and S. F. McCormick: *First-order system least squares for the Navier–Stokes equations.* In Melson, N. D., T. A. Manteuffel, S. F. McCormick and C. C. Douglas (editors): *Seventh Copper Mountain Conference on Multigrid Methods*, volume CP 3339, pages 41–55, Hampton, VA, 1996. NASA.

[14] Bochev, P., Z. Cai, T. A. Manteuffel and S. F. McCormick: *Analysis of Velocity-Flux First-Order System Least-Squares Principles for the Navier-Stokes Equations: Part I.* SIAM J. Numer. Anal., 35(3):990–1009, 1998.

[15] Bochev, P. and M. Gunzburger: *Analysis of least-squares finite element methods for the Stokes equations.* Math. Comp., 63:479–506, 1994.

[16] Bochev, P. and M. Gunzburger: *Analysis of weighted least-squares finite element method for the Navier-Stokes equations.* In *Proceedings of the 14th IMACS World Congress*, pages 584–587, Georgia Tech, Atlanta, 1994.

[17] Bochev, P. and M. Gunzburger: *Least-squares for the velocity-pressure-stress formulation of the Stokes equations.* Comput. Methods Appl. Mech. Engrg., 126:267–287, 1995.

[18] Bochev, Pavel, 2004. e-mail correspondence.

[19] Bochev, Pavel and Max Gunzburger: *Finite Element Methods of Least-Squares Type.* SIAM Rev., 40(4):789–837, 1998.

[20] Bochev, Pavel, Thomas A. Manteuffel and Stephen F. McCormick: *Analysis of Velocity-Flux Least-Squares Principles for the Navier–Stokes Equations: Part II.* SIAM J. Numer. Anal., 36(4):1125–1144, 1999.

[21] Bochev, Pavel B.: *Least Squares Finite Element Methods for the Stokes and Navier-Stokes Equations.* PhD thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, July 1994.

[22] Bolton, P., J. Stratakis and R. W. Thatcher: *Mass conservation in least squares methods for Stokes flow.* Technical Report, Department of Mathematics,

University of Manchester Institute of Science and Technology, Manchester, UK, 2001. http://www.ma.umist.ac.uk/rwt/research.htm.

[23] Braess, D.: *Finite Elements.* Cambridge University Press, Cambridge, U.K., 1997.

[24] Bramble, James H., Raytcho D. Lazarov and Joseph E. Pasciak: *A least-squares approach based on a discrete minus one inner product for first order systems.* Math. Comp., 219:935–956, 1997.

[25] Bramble, James H., Raytcho D. Lazarov and Joseph E. Pasciak: *Least-Squares methods for linear elasticity based on a discrete minus one inner product.* Comput. Methods Appl. Mech. Engrg., 152:520–543, 2001.

[26] Bramble, James H. and Joachim A. Nitsche: *A Generalized Ritz-Least-Squares Method for Dirichlet Problems.* SIAM J. Numer. Anal., 10(1):81–93, 1973.

[27] Brenner, Susanne C. and L. Ridgway Scott: *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, Berlin, 1994.

[28] Brezzi, F. and R. S. Falk: *Stability of higher-order Hood-Taylor methods.* SIAM J. Numer. Anal., 28(3):581–590, 1991.

[29] Brezzi, F. and L. D. Marini: *A three-field domain decomposition method.* In Quarteroni, A. (editor): *Domain Decomposition Methods in Science and Engineering*, pages 27–34. American Mathematical Society, Providence, RI, 1994.

[30] Brezzi, Franco and Michel Fortin: *Mixed and hybrid finite element methods.* Number 15 in *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, 1991.

[31] Brown, P.N., B.Lee and T.A. Manteuffel: *A Moment-Parity Multigrid Preconditioner for the first-order system least-squares formulation of the Boltzmann Transport Equation.* SIAM J. Sci. Statist. Comput., 25(2):513–533, 2003.

[32] Burenkov, Victor I.: *Sobolev Spaces on Domains*, volume 137 of *Teubner-Texte zur Mathematik.* Teubner, Stuttgart, Leipzig, 1998.

[33] Burrell, Laura L., Li Q. Tang and Tate T.H. Tsang: *On a Least-Squares Finite Element Method for Advective Transport in Air Pollution Modelling.* Atmospheric Environment, 29(12):1425–1439, 1995.

[34] Cai, Z., T. A. Manteuffel and S. F. McCormick: *First-order system least squares for velocity-vorticity-pressure form of the Stokes equations, with applications to linear elasticity.* In Melson, N. D., T. A. Manteuffel, S. F. Mc-

Cormick and C. C. Douglas (editors): *Seventh Copper Mountain Conference on Multigrid Methods*, volume CP 3339, pages 123–132, Hampton, VA, 1996. NASA.

[35] Cai, Z., T. A. Manteuffel and S. F. McCormick: *First-order system least squares for the Stokes equations, with application to linear elasticity.* SIAM J. Numer. Anal., 34(5):1727–1741, 1997.

[36] Cai, Z., T.A. Manteuffel, S.F. McCormick and J. Ruge: *First-Order System LL* (FOSLL*): Scalar Elliptic Partial Differential Equations.* SIAM J. Numer. Anal., 39:1418–1445, 2001.

[37] Cai, Z. and Gerhard Starke: *First-order system least squares for the stress-displacement formulation: Linear elasticity.* SIAM J. Numer. Anal., 41:715–730, 2003.

[38] Cai, Z. and Gerhard Starke: *Least-Squares Methods for Linear Elasticity.* SIAM J. Numer. Anal., 42:826–842, 2004.

[39] Cai, Zhiqiang: *Least Squares for the Perturbed Stokes equations and the Reissner-Mindlin plate.* SIAM J. Numer. Anal., 38(5):1561–1581, 2000.

[40] Cai, Zhiqiang, Johannes Korsawe and Gerhard Starke: *An adaptive least squares mixed finite element method for the stress-displacement formulation of linear elasticity.* Numer. Methods Part. Diff. Eqs., 21:132–148, 2005.

[41] Cai, Zhiqiang, Barry Lee and Ping Wang: *Least-Squares Methods for Incompressible Newtonian Fluid Flow: Linear Stationary Problem.* SIAM J. Numer. Anal., 42(2):843–859, 2004.

[42] Cai, Zhiqiang, Chang-Ock Lee, Thomas A. Manteuffel and S.F. McCormick: *First-order System Least-Squares for linear elasticity: Further Results.* SIAM J. Sci. Statist. Comput., 21(5):1728–1739, 2000.

[43] Cai, Zhiqiang, Chang-Ock Lee, Thomas A. Manteuffel and S.F. McCormick: *First-order System Least-Squares for linear elasticity: Numerical Results.* SIAM J. Sci. Statist. Comput., 21(5):1706–1727, 2000.

[44] Cai, Zhiqiang, Thomas A. Manteuffel, S.F. McCormick and Seymour Parter: *First-Order System Least Squares (FOSLS) for Planar Linear Elasticity: Pure Traction Problem.* SIAM J. Numer. Anal., 35(1):320–335, 1998.

[45] Cai, Zhiqiang, Thomas A. Manteuffel and Stephen F. McCormick: *First-order system least squares fort velocity-vorticity-pressure form of the Stokes equations, with application to linear elasticity.* Elect. Trans. Numer. Anal., 3:150–159, 1995.

[46] Cebral, J.R. and R. Löhner: *Conservative load projection and tracking for fluid-structure problems*. AIAA J., 35(4):687–692, 1997.

[47] Chang, C.L.: *A mixed finite element method for Stokes problem: an acceleration-pressure formulation*. Appl. Math. Comput., 36:135–146, 1990.

[48] Chang, C. L. and B. N. Jiang: *An error analysis of least-squares finite element method of velocity-pressure-vorticity formulation for the Stokes problem*. Comput. Methods Appl. Mech. Engrg., 84:247–255, 1990.

[49] Chang, C.L. and John L. Nelson: *Least-Squares Finite Element Method for the Stokes Problem with Zero Residual Mass Conservation*. SIAM J. Numer. Anal., 34(2):480–489, 1997.

[50] Chen, T. F.: *Semidiscrete least squares methods for linear convection-diffusion problems*. Comput. Math. Appl., 24:29–44, 1992.

[51] Ciarlet, P. G.: *Mathematical Elasticity Volume I: Three-Dimensional Elasticity*. North-Holland, Amsterdam, 1988.

[52] Cockburn, B., G. E. Karniadakis and C.-W. Shu (editors): *Discontinuous Galerkin Methods: Theory, Computation and Applications*. Number 11 in *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2000.

[53] Codd, A., T. Manteuffel, S. McCormick and J. W. Ruge: *Multilevel first-order system least squares for elliptic grid generation*. SIAM J. Numer. Anal., 41(6):2210–2232, 2003.

[54] Codd, A. L.: *Elasticity-Fluid Coupled Systems and Elliptic Grid Generation (EGG) based on First-Order System Least Squares (FOSLS)*. PhD thesis, University of Colorado, Department of Applied Mathematics, 2001.

[55] Codd, A.L., T.A. Manteuffel and S.F. McCormick: *Multilevel First-Order system least squares for nonlinear elliptic partial differential equations*. SIAM J. Numer. Anal., 41(6):2197–2209, 2003.

[56] Curnier, A.: *Computational Methods in Solid Mechanics*. Kluwer Academic Publishers Group, Dordrecht, The Netherlands, 1994.

[57] Dautray, Robert and Jacques-Louis Lions: *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 5. Springer-Verlag, Berlin, 1992.

[58] Deang, Jennifer M. and Max D. Gunzburger: *Issues related to Least-Squares Finite Element Methods for the Stokes equations*. SIAM J. Sci. Statist. Comput., 20(3):878–906, 1998.

[59] Donea, J. and L. Quartapelle: *An introduction to finite element methods for transient advection problems.* Internat. J. Numer. Methods Engrg., 95:169–203, 1992.

[60] Donea, Jean and Antonio Huerta: *Finite Element Methods for Flow Problems.* John Wiley & Sons, New York, London, Sydney, 2003.

[61] Du, Q., M.D. Gunzburger, L.S. Hou and J. Lee: *Analysis of a Linear Fluid-Structure Interaction Problem.* Disc. Cont. Dyn. Syst., (9):633–650, 2003.

[62] Du, Q., M.D. Gunzburger, L.S. Hou and J. Lee: *Semidiscrete Finite Element Approximations of a Linear Fluid-Structure Interaction Problem.* SIAM J. Numer. Anal., 42(1):1–29, 2004.

[63] Edgar, Nathan B. and Karan S. Surana: *p-Version least squares finite element formulation for axisymmetric incompressible non-Newtonian fluid flow.* Comput. Methods Appl. Mech. Engrg., 113:271–300, 1994.

[64] Farhat, C., P. Geuzaine and C. Grandmont: *The discrete geometric conservation law and the nonlinear stability of ALE schemes for the solution of flow problems on moving grids.* J. Comput. Phys., 174, 2001.

[65] Farhat, Charbel and Francois-Xavier Roux: *A method of finite element tearing and interconnecting and its parallel solution algorithm.* Internat. J. Numer. Methods Engrg., 32:1205–1227, 1991.

[66] Ferziger, Joel. H. and Milovan Perić: *Computational Methods for Fluid Dynamics.* Springer-Verlag, Berlin, 1999. second edition.

[67] Fiard, J. M., T. A. Manteuffel and S. F. McCormick: *First-Order System Least Squares (FOSLS) for Convection-Diffusion Problems: Numerical Results.* SIAM J. Sci. Statist. Comput., 19(6):1958–1979, 1998.

[68] Fortin, M. and M. Soulie: *A non-conforming piecewise quadratic finite element on triangles.* Internat. J. Numer. Methods Engrg., 19:505–520, 1983.

[69] Geller, S., M. Krafczyk, J. Tölke, S. Turek and J. Hron: *Benchmark computations based on Lattice-Boltzmann, Finite Element and Finite Volume Methods for laminar Flows.* Ergebnisberichte des Fachbereichs Mathematik 274, Universität Dortmund, 2004.

[70] Girault, Vivette and Pierre-Arnaud Raviart: *Finite Element Methods for Navier-Stokes Equations.* Number 5 in *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, 1986.

[71] Glowinski, R., B. Mantel, J. Periaux, P. Perrier and O. Pironneau: *On an Efficient New Preconditioned Conjugate Gradient Method. Application to the In-core Solution of the Navier-Stokes Equations Via Non-linear Least-Squares and Finite Element Methods.* Finite Elements in Fluids, 4:365–401, 1982.

[72] Gresho, P.M. and R.L. Sani: *Incompressible Flow and the Finite Element Method.* John Wiley and Sons, 2000.

[73] Hackbusch, W.: *Multi-Grid Methods and Applications.* Number 4 in *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, 1985.

[74] Haupt, Peter: *Continuum Mechanics and Theory of Materials.* Springer-Verlag, Berlin, 2000.

[75] Heys, Jeffrey J., Curt G. DeGroff, Thomas A. Manteuffel, Stephen F. McCormick and Henry Tufo: *Modelling 3-D Compliant Blood Flow With FOSL.* Biomed. Sci. Instr., 2004. to appear.

[76] Heys, Jeffrey J., Curt G. DeGroff, Wendy W. Orlando, Thomas A. Manteuffel and Stephen F. McCormick: *First-Order System Least Squares for Elastohydrodynamics with Application to Flow in Compliant Blood Vessels.* Biomed. Sci. Instr., 38:277–282, 2002.

[77] Heys, Jeffrey J., Thomas A. Manteuffel, Stephen F. McCormick and J. Ruge: *First-order system least squares for coupled fluid-elasticity problems.* J. Comput. Phys., 195(2):560–575, 2004.

[78] Hrenikoff, A.: *Solution of problems in elasticity by the framework method.* J. Appl. Mech., A8:169–175, 1941.

[79] Hübner, B., E. Walhorn and D. Dinkler: *A monolithic approach for fluid-structure interaction with space-time finite elements.* Comput. Methods Appl. Mech. Engrg., 2004.

[80] Hübner, Björn: *Simultane Analyse von Bauwerk-Wind-Wechselwirkungen.* PhD thesis, TU Braunschweig, 2002.

[81] Hughes, T., L. Franca and M. Ballestra: *A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accomodating equal-order interpolations.* Comput. Methods Appl. Mech. Engrg., 59:85–99, 1986.

[82] Jiang, B. N.: *Non-oscillatory and non-diffusive solution of convection problems by the iteratively reweighted least-squares finite element method.* J. Comput. Phys., 105:108–121, 1993.

[83] Jiang, B. N.: *The Least-Squares Finite Element Method*. Springer-Verlag, Berlin, 1998.

[84] Jiang, B. N. and C. L. Chang: *Least Squares finite elements for Stokes problem*. Comput. Methods Appl. Mech. Engrg., 78:297–311, 1990.

[85] Jiang, B. N., T. L. Lin and L. A. Povinelli: *Large-scale computation of incompressible viscous flow by least-squares finite element method*. Comput. Methods Appl. Mech. Engrg., 114:213–231, 1994.

[86] Kayser-Herold, Oliver: *Least-Squares FEM, Literature Review*. Informatik-Bericht 2005-05, TU Braunschweig, Institut für Wissenschaftliches Rechnen, 2005. `http://opus.tu-bs.de/opus/volltexte/2005/758/`.

[87] Kayser-Herold, Oliver and Hermann G. Matthies: *Space-time adaptive solution of fluid-structure interaction problems*. In Bathe, Klaus-Jürgen (editor): *MIT 2nd Conference on Computational Fluid and Solid Mechanics*, pages 1000–1004. Elsevier Sciences Ltd., 2003.

[88] Kim, Sang Dong, Thomas A. Manteuffel and Stephen F. McCormick: *First-Order System Least Squares (FOSLS) for Spatial Linear Elasticity: Pure Traction*. SIAM J. Numer. Anal., 38(5):1454–1482, 2000.

[89] Lazarov, R. D., L. Tobiska and P. S. Vassilevski: *Streamline Diffusions Least-Squares Mixed Finite Element Methods for Convection-Diffusion problems*. In *Forschungsberichte der Fakultät für Mathematik*, number 29. Otto-von-Guericke-Universität Magdeburg, 1997. Preprint.

[90] Lee, Chang-Ock: *Multigrid Methods for the Pure Traction Problem of Linear Elasticity: Mixed Formulation*. SIAM J. Numer. Anal., 35(1):121–145, 1998.

[91] Lee, Sung-Ho, Sung-Kie Youn, Jeoung-Heum Yeon and Bo-Nan Jiang: *A Study on the Fluid-Structure Interaction using LSFEM*. Technical Report, Department of Mechanical Engineering, KAIST, 2000.

[92] Maischak, M. and E.P. Stephan: *A Least Squares coupling method with finite elements and boundary elements for transmission problems*. Comput. Math. Appl., 48:995–1016, 2004.

[93] Majidi, Mohammad and Gerhard Starke: *Least-Squares Galerkin Methods for Parabolic Problems I: Semidiscretisation in Time*. SIAM J. Numer. Anal., 39(4):1302–1323, 2001.

[94] Majidi, Mohammad and Gerhard Starke: *Least-Squares Galerkin Methods for Parabolic Problems II: The Fully Discrete Case and Adaptive Algorithms*. SIAM J. Numer. Anal., 39(5):1648–1666, 2002.

[95] Manteuffel, Thomas A. and Klaus J. Ressel: *Least-Squares finite-element solution of the Neutron transport equation in diffusive regimes.* SIAM J. Numer. Anal., 35(2):806–835, 1998.

[96] Matthies, Hermann G. and Jan Steindorf: *Partioned strong coupling algorithms for fluid-structure interaction.* Comput. & Structures, 81:805–812, 2003.

[97] Matthies, Hermann G. and Jan Steindorf: *Strong Coupling Methods.* In Wendland, Wolfgang and Messoud Efendiev (editors): *Analysis and Simulation of Multifield Problems*, volume 12 of *Lecture Notes in Applied and Computational Mechanics*, pages 13–36. Springer-Verlag, Berlin, 2003.

[98] McCormick, S. F.: *FOSLoSophy, Informal discussion of some advantages and disadvantages of First-Order System Least Squares (FOSLS).* Technical Report, University of Colorado, 2000. `http://amath.colorado.edu/pub/fosls/FOSLoSophy.ps`.

[99] McHenry, D.: *A lattice analogy for the solution of plane stress problems.* J. Inst. Civ. Eng., 21:59–82, 1943.

[100] Mouro, J. and P. LeTallec: *Fluid structure interaction with large structural displacements.* Comput. Methods Appl. Mech. Engrg., 190:3039–3067, 2001.

[101] Nguyen, H. and J. Reynen: *A space-time least-squares finite element scheme for advection-diffusion equations.* Comput. Methods Appl. Mech. Engrg., 42:331–342, 1984.

[102] Olson, Luke Nathan: *Multilevel Least-Squares Finite Element Methods for Hyperbolic Partial Differential Equations.* PhD thesis, University of Colorado, Department of Applied Mathematics, 2003.

[103] Park, K. C. and C. A. Felippa: *Recent development in coupled field analysis methods.* In Lewis, R. W., P. Bettess and E. Hinton (editors): *Numerical Methods in Coupled Systems*, pages 327–352. John Wiley & Sons, New York, London, Sydney, 1984.

[104] Park, N. S. and J. A. Liggett: *Taylor-least-squares finite element for two-dimensional advection-dominated unsteady advection-diffusion problems.* Internat. J. Numer. Methods Fluids, 11(1):21–38, 1990.

[105] Pedersen, Michael: *Functional Analysis in Applied Mathematics and Engineering.* Studies in advanced mathematics. Chapman & Hall/CRC, Boca Raton, 2000.

[106] Piperno, S.: *Explicit/implicit fluid/structure staggered procedures with a structural predictor and fluid subcycling for 2D inviscid aeroelastic simulations.* Internat. J. Numer. Methods Fluids, 25:1207–1226, 1997.

[107] Piperno, Serge: *Numerical simulation of aeroelastic instabilities of elementary bridge decks.* Technical Report 3549, Institut National de Recherche en Informatique et en Automatique (INRIA), Septembre 1998.

[108] Quarteroni, Alfio, Massimiliano Tuveri and Alessandro Veneziani: *Computational vascular fluid dynamics: problems, models and methods.* Comput. Visual. Sci., 2(4):163–197, 2000.

[109] Quarteroni, Alfio and Alberto Valli: *Numerical Approximation of Partial Differential Equations.* Springer-Verlag, Berlin, 1997.

[110] Quarteroni, Alfio and Alberto Valli: *Domain Decomposition Methods for Partial Differential Equations.* Oxford University Press, New York, Oxford, 1999.

[111] Rachford, H. H. and M. F. Wheeler: *An H^{-1} Galerkin Procedure for the Two-Point Boundary-Value Problem.* In Boor, C. de (editor): *Mathematical Aspects of Finite Element Methods in Partial Differential Equations*, pages 353–382. Academic Press, New York, 1975.

[112] Reese, S.: *Theorie und Numerik des Stabilitätsverhaltens hyperelastischer Festkörper.* PhD thesis, Technische Hochschule Darmstadt, 1994.

[113] Salonen, E. and J. Freund: *Weighting in the Least Squares Finite Element Method.* In Mang, H.A., F.G. Rammerstorfer and J. Eberhardsteiner (editors): *Proceedings of the Fifth World Congress on Computational Mechanics (WCCM V)*, Vienna, Austria, July 7-12 2002. Vienna University of Technology, Austria.

[114] Schäfer, M. and S. Turek: *Benchmark computations of laminar flow around a cylinder.* In Hirschel, E.H. (editor): *Flow Simulations with High-Performance Computers II*, volume 52, pages 547–566. Vieweg, Braunschweig, 1996.

[115] Schwarz, H.R.: *Numerische Mathematik.* Teubner, Stuttgart, Leipzig, 1997.

[116] Sohr, Hermann: *The Navier-Stokes Equations, An Elementary Functional Analytic Approach.* Birkhäuser Verlag, Basel, 2001.

[117] Starke, Gerhard: *Multilevel boundary functionals for least-squares mixed finite element methods.* SIAM J. Numer. Anal., 36:1065–1077, 1999.

[118] Starke, Gerhard: *A First order System Least Squares Finite Element Method for the Shallow Water Equations.* SIAM J. Numer. Anal., 42:2387–2407, 2005.

[119] Steindorf, Jan: *Partitionierte Verfahren für Probleme der Fluid-Struktur Wechselwirkung*. PhD thesis, TU Braunschweig, 2002.

[120] Tang, Li Q., Tiwu Cheng and Tate T.H. Tsang: *Transient Solutions for Three-Dimensional Lid-Driven Cavity Flows by a Least-Squares Finite Element Method*. Internat. J. Numer. Methods Fluids, 21:413–432, 1995.

[121] Tang, Li Q. and Tate T.H. Tsang: *A Least-Squares Finite Element Method for Time-Dependent Incompressible Flows with Thermal Convection*. Internat. J. Numer. Methods Fluids, 17:271–289, 1993.

[122] Tang, Li Q., Jamie L. Wright and Tate T.H. Tsang: *Simulations of 2D and 3D Thermocapillary Flows by a Least-Squares Finite Element Method*. Internat. J. Numer. Methods Fluids, 28:983–1007, 1998.

[123] Temam, Roger: *Navier-Stokes Equations and Nonlinear Functional Analysis*. SIAM Publ., Philadelphia, 1995. 2nd edition.

[124] Thatcher, R. W.: *A least-squares method for Stokes flow based on stress and stream functions*. Technical Report, Department of Mathematics, University of Manchester Institute of Science and Technology, Manchester, UK, 1997. `http://www.ma.umist.ac.uk/rwt/research.htm`.

[125] Thomeée, Vidar: *Galerkin Finite Element Methods for Parabolic Problems*. Springer Series in Computation Mathematics. Springer-Verlag, Berlin, 1997.

[126] Turek, Stefan: *Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approach*, volume 6 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, 1999.

[127] Walhorn, Elmar: *Ein simultanes Berechnungsverfahren für Fluid-Struktur-Wechselwirkungen mit finiten Raum-Zeit-Elementen*. PhD thesis, TU Braunschweig, 2002.

[128] Wall, W.A.: *Fluid-Struktur Interaktion mit stabilisierten Finiten Elementen*. PhD thesis, Universität Stuttgart, 1999.

[129] Wendland, Wolfgang L.: *Elliptic Systems in the Plane*. Pitman, Boston, 1979.

[130] Winterscheidt, Daniel and Karan S. Surana: *p-Version Least Squares Finite Element Formulation for Two-Dimensional, Incompressible Fluid Flow*. Internat. J. Numer. Methods Fluids, 18:43–69, 1994.

[131] Wohlmuth, Barbara I.: *Discretization Methods and Iterative Solvers Based on Domain Decomposition*. Number 17 in *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2001.

[132] Wood, W.L. and R.W. Lewis: *A comparision of time marching schemes for the transient heat conduction equation*. Internat. J. Numer. Methods Engrg., 9, 1975.

[133] Wu, J. and B. N. Jiang: *A least-squares finite element method for electromagnetic scattering problems*. In Reddy, J.N. (editor): *Proceedings of the Third US National Congress on Computational Mechanics*, (Texas A & M University, College Station), 1995.

[134] Wu, J. and B. N. Jiang: *A time-accurate least-squares finite-element method for incompressible inviscid and viscous flows*. Technical Report, ICOMP, 1995. (unpublished).

[135] Yang, Suh-Yuh and Ching L. Chang: *Analysis of a Two-stage Least-squares Finite Element Method for the Planar Elasticity Problem*. Math. Methods Appl. Sci., 22:713–732, 1999.

[136] Zienkiewicz, O.C. and R.W. Lewis: *An analysis of various time stepping schemes for initial value problems*. Earthquake Engrg. Struct. Dyn., 1:407–408, 1973.

[137] Zienkiewicz, O.C. and R.L. Taylor: *The Finite Element Method*. Butterworth Heinemann, Oxford, 5th edition, 2000.